



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 164528

**TO: Anish Gupta**  
**Art Unit: 1654**  
**Location: REM/3C15/3C18**  
**Serial Number: 10/083768**

**Friday, September 02, 2005**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### Search Notes

#### Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

This Page Blank (uspto)



FOR OFFICIAL USE ONLY

CP 66

ACCESS DB # 164518  
PLEASE PRINT CLEARLY

Scientific and Technical Information Center

SEARCH REQUEST FORM

Requester's Full Name: Anush Gupta Examiner #: 73121 Date: 8/31/01  
Art Unit: 1654 Phone Number: 2-0965 Serial Number: 10/083,768  
Location (Bldg/Room#): Rosen/608 (Mailbox #): \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK  
\*\*\*\*\*

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: Peptide & Compounds that bind to a Receptor  
Inventors (please provide full names): William Dawes, Ronald Barrett, Steven Cwikla,  
David Duffin, Christian Gatea, Steven Hazelden, Larry Madzickis, Peter Schatz,  
Earliest Priority Date: 6/7/96 Christopher Waghorn, Nicholas Wrighton

Search Topic:

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

nej Please search seq Id No. 6-13

6 aa 18  
7 19  
8 19  
9 18  
10 18  
11 19  
12 14  
13 aa 14

Complete  
all info  
Int.

STAFF USE ONLY

Searcher: Beverly C-25-2-8 Type of Search \_\_\_\_\_ NA Sequence (#)

Searcher Phone #: \_\_\_\_\_ AA Sequence (#)

Searcher Location: \_\_\_\_\_ Structure (#)

Date Searcher Picked Up: \_\_\_\_\_ Bibliographic

Date Completed: \_\_\_\_\_ Litigation

Searcher Prep & Review Time: \_\_\_\_\_ Fulltext

Online Time: \_\_\_\_\_ Other

Vendors and cost where applicable

\_\_\_\_\_ STN \_\_\_\_\_ Dialog

\_\_\_\_\_ Questel/Orbit \_\_\_\_\_ Lexis/Nexis

\_\_\_\_\_ Westlaw \_\_\_\_\_ WWW/Internet

\_\_\_\_\_ Lab house sequence systems CGN

\_\_\_\_\_ Commercial \_\_\_\_\_ Oligomer \_\_\_\_\_ Score/Length

\_\_\_\_\_ Interference \_\_\_\_\_ SPDI \_\_\_\_\_ Encode/Transl

\_\_\_\_\_ Other (specify)

**This Page Blank (uspto)**



99 56 51.4 13 5 ABB72901  
100 56 51.4 13 7 ADJ73054

Abb72901 TPO mimet  
Adj73054 TPO mimet

## ALIGNMENTS

RESULT 1  
AAW09456  
ID AAW09456 standard; protein; 18 AA.

XX AAW09456;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound peptide.

KM Haematology; thrombocytopenia; TPO; TR; proliferation;

KW bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 1.18  
/note= "Preferably linkages are selected from: -  
CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -C(O)NR6  
; -NH(C(O)NH; where R is hydrogen or lower alkyl and R6 is  
lower alkyl"

FT Modified-site

1  
/note= "Preferably N-terminus is selected from: -NRR1; -  
NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NH(C(O)NR; succinimide;  
benzyloxycarbonyl-NH; benzyloxycarbonyl-NH with 1-3  
substitutions on the phenyl ring selected from lower  
alkyl, lower alkoxy, chloro, bromo; where R and R1 are  
independently selected from hydrogen and lower alkyl"

FT Modified-site

18  
/note= "Preferably C-terminus is -C(O)R2 where R2 is  
selected from hydroxy, lower alkoxy, and -NR3R4, where R3  
and R4 are independently selected from hydrogen and lower  
alkyl, and where the nitrogen atom of the -NR3R4 group  
can optionally be the amine group of the N-terminus of  
the peptide forming a cyclic peptide"

XX W09640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

XX mimetic(s) - useful in treatment of haematological disorders, esp.

XX thrombocytopenia resulting from chemotherapy, etc.

CC treating patients suffering from haematological disorders and  
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
CC marrow transfusions. The peptide may also be used to maintain the  
CC proliferation and growth of TPO-dependent cell lines and for use in  
CC biological research, for detecting TPO receptors on living cells  
XX  
SQ Sequence 18 AA;  
Query Match 100.0%; Score 109; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCADGPTLRWISFCGG 18  
1 |||||  
DB 1 GGCADGPTLRWISFCGG 18  
RESULT 2  
AAW33023  
ID AAW33023 standard; peptide; 18 AA.  
XX AAW33023;  
AC AAW33023;  
XX  
DT 11-MAR-1998 (first entry)  
DE Thrombopoietin receptor binding peptide.  
XX  
XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
XX haematological disorder; thrombocytopenia; chemotherapy;  
XX radiation therapy; bone marrow transfusion; diagnosis;  
XX signal transduction; receptor activation; cell culture.  
OS Synthetic.  
XX  
XX W09640750-A1.  
PN  
PD 19-DEC-1996.  
PF 07-JUN-1996; 96WO-US009623.  
XX  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
XX PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX WPI; 1997-052226/05.  
XX  
XX Peptides and peptide mimetics which bind to and activate the  
XX Thrombopoietin receptor - useful in treatment of haematological  
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
PS Claim 19; Page 89; 106pp; English.  
XX  
XX The present peptide binds the thrombopoietin receptor (TR), has a  
XX molecular weight of less than 8000 Da and a TR binding affinity as  
XX expressed by an IC50 of no more than about 100 microm. It can be used to  
XX treat disorders which are susceptible to treatment with a thrombopoietin  
XX agonist, preferably haematological disorders and thrombocytopenia  
XX resulting from chemotherapy, radiation therapy or bone marrow  
XX transfusions. It can also be used diagnostically, e.g. to investigate the  
XX mechanism of thrombopoietin signal transduction and receptor activation,  
XX or to maintain the proliferation and growth of thrombopoietin dependent  
XX cell lines  
XX  
SQ Sequence 18 AA;  
Query Match 100.0%; Score 109; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 ID |||||  
 DB 1 GGCADGPTLRWISFCGG 18

## RESULT 3

AA17020  
 ID AAB17020 standard; peptide; 18 AA.

AC AAB17020;  
 XX

DT 31-OCT-2000 (first entry)

DE TPO-mimetic peptide sequence SEQ ID NO:76.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytotoxic; antitumor; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTU44; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.

XX Synthetic.

PN WO200024782-A2.

PD 04-MAY-2000.

PF 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Fette U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

PT Novel composition of matter comprising an Fc domain and pharmacologically  
 active peptides, useful for treating cancer and autoimmune diseases.

PS Claim 19; Page 220; 608pp; English.

XX The present invention describes composition of matter (1) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (1) is:  
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-F1, -(L1)-c-P1-(L2)-d-P2, -(L1)-c-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antitumor, and host  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AA69443 to AA69526 and AA61655 to  
 CC AA61803 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention

XX Sequence 18 AA;

Query Match 100.0%; Score 109; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 DB 1 GGCADGPTLRWISFCGG 18

RESULT 4  
 ID AAU25820  
 AAU25820 standard; peptide; 18 AA.

AC AAU25820;  
 XX

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #6.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ;

XX Balasubramanian P, Wagstrom CR, Hendren RM, Deprience RB, Poddaturi S;

XX Yin Q;

XX WPI; 2001-564142/63.

XX Disclosure; Col 65-66; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 18 AA;

Query Match 100.0%; Score 109; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 DB 1 GGCADGPTLRWISFCGG 18

Db 1 GGCADGPTLRWISFCGG 18

RESULT 5  
ABR72906

ID ABR72906 standard; peptide; 18 AA.

XX ABR72906;

AC ABR72906;

DT 05-APR-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:76.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
KM erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
KM TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
KM TPO mimetic peptide; EPO mimetic peptide; EME; VEGF antagonist;  
KM MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
KM cyclostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
KM antianemic; anorectic; antiinfertility; haemostatic; dermatological;  
KM neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
KM cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
KM sleep disorder; neurological degenerative disease; anaemia;  
KM thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
KM Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

OS WO200183525-A2.

XX 08-NOV-2001.

PD 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

PR (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

DR Novel vehicle-peptide molecule or its multimers useful for treating  
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 44; 17pp; English.

PS The present invention describes a vehicle-peptide molecule (I) or its  
XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and  
CC neuroprotective activities. (I) can be used as a therapeutic or  
CC prophylactic agent as well as for screening purposes. (I) is useful for  
CC diagnosing diseases characterised by dysfunction of their associated  
CC protein of interest, for identifying normal or abnormal proteins of  
CC interest, as a part of diagnostic kit to detect the presence of their  
CC proteins of interest in a biological sample. Additionally, (I) is useful  
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
CC mimetic compounds are useful for treating disorders characterised by low  
CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
CC compounds are useful for treating conditions that involve an existing  
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABR72403 to ABR73426 and ABR35695 to ABR35777  
CC represent amino acid and nucleic acid sequences used in the  
XX exemplification of the present invention

XX Sequence 18 AA;

SO

Query Match 100.0%; Score 109; DB 5; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
|||||

Db 1 GGCADGPTLRWISFCGG 18

RESULT 6  
ADJ73058

ID ADJ73058 standard; peptide; 18 AA.

XX ADJ73058;

AC ADJ73058;

DT 06-MAY-2004 (first entry)

DE TPO mimetic peptide sequence SeqID 512.

XX mimetic; CDR mimeticbody; gene therapy; transgenic; immune;  
KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
KM immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
KM TPO.

XX Synthetic.

OS WO2003084477-A2.

XX 16-OCT-2003.

PD 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

PR (CENZ) CENTOCOR INC.

PA Heavner GA, Knight DM, Scallion BJ, Ghayab J;

XX WPI; 2003-804237/75.

DR New CDR mimeticbody comprising a portion of a heavy or light chain  
PT variable region comprising human framework or ligand binding region,  
PT useful for preparing a composition for treating e.g., immune,  
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 512; 97pp; English.

PS This invention relates to novel mammalian CDR mimeticbodies, specific  
CC portions or variants thereof. Specifically, it refers to an antibody  
CC fragment where a protein has been inserted into, or replaces a portion  
CC of, one or more CDR regions, such that each CDR mimeticbody comprises at  
CC least one portion of a heavy chain or light chain variable region, which  
CC itself comprises at least one human framework region and at least one  
CC ligand binding region (LBR). The present invention describes human  
CC mimeticbodies, including modified immunoglobulins and cleavage products  
CC that can be useful in gene therapy and the generation of transgenic  
CC plants and animals. Furthermore, the CDR mimeticbody is useful for  
CC preparing compositions for modulating, treating or reducing the symptoms  
CC of immune, cardiovascular, infectious, malignant and/or neurologic  
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
CC peptide sequence is a TPO mimetic peptide sequence used to make a  
CC mimeticbody of the invention.

XX Sequence 18 AA;

SO Query Match 100.0%; Score 109; DB 7; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
|||||

Db 1 GGCADGPTLRMISFCG 18

RESULT 7  
ADJ52693 ID ADJ52693 standard; peptide; 18 AA.  
XX  
XX ADJ52693;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX CH1 deleted mimetibody-related peptide SegID512.  
XX  
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiac;  
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
KW arhythmia; hypertension; heart failure; neurodegenerative;  
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
KW cancerous condition; infectious disease; bacterial infection;  
KW viral infection; fungal infection.  
XX  
XX Unidentified.  
OS Synthetic.  
XX  
XX WO2004002417-A2.  
XX  
XX 08-JAN-2004.  
XX  
XX 27-JUN-2003; 2003WO-US020347.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX  
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
PI Kutolowski KA;  
PI  
XX WPI; 2004-082870/08.  
XX  
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
PT diseases.  
XX  
XX Claim 2; SEQ ID NO 512; 123pp; English.  
XX  
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an immunosuppressive,  
CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,  
CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
CC is useful for diagnosing or treating a disease condition in a cell,  
CC tissue, organ or animal, specifically for modulating, treating,  
CC alleviating, preventing, the incidence or reducing the symptoms of an  
CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
CC conditions, or infectious diseases (for example bacterial, viral or  
CC fungal infection). The present sequence is that of a peptide which may be  
CC used during the creation of a mimetibody of the invention.  
XX  
XX Sequence 18 AA;  
SQ

Query March 100.0%; Score 109; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGCADGPTLRMISFCG 18  
1 GGCADGPTLRMISFCG 18

RESULT 8  
ADJ51654 ID ADJ51654 standard; peptide; 18 AA.  
XX  
XX ADJ51654;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX CH1 deleted mimetibody-related peptide SegID512.  
XX  
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
KW obstetric disorder; haematologic disorder; immunologic disorder;  
KW allergic disorder; infectious disorder; musculoskeletal disorder;  
KW oncological disorder; neurological disorder; nutritional disorder;  
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
KW renal disorder; pulmonary disorder.  
XX  
XX Unidentified.  
OS Synthetic.  
XX  
XX WO2004002424-A2.  
XX  
XX 08-JAN-2004.  
XX  
XX 30-JUN-2003; 2003WO-US020495.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX 19-SEP-2002; 2002US-0412144P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX  
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
PI Kutolowski KA;  
PI  
XX WPI; 2004-082872/08.  
XX  
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for  
PT diagnosing, preventing or treating cardiovascular, dermatologic,  
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
PT nutritional disorders.  
XX  
XX Claim 15; SEQ ID NO 512; 123pp; English.  
XX  
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an osteopathic,  
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
CC gastroenteric-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or  
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
CC modulator or cytokine-agonist. The methods and compositions of the  
CC present invention are useful for the diagnosis, prevention and/or  
CC treatment of diseases or conditions associated with aberrant expression  
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,  
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
CC obstetric, haematologic, immunologic, allergic, infectious,  
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
CC pediatric, psychiatric, renal or pulmonary disorders. The present  
CC sequence is that of a peptide which may be used during the creation of a  
CC mimetibody of the invention.  
XX  
XX Sequence 18 AA;  
SQ





Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCADGPTLREWISFCGG 18  
 |||||  
 Db 1 GGCADGPTLREWISFCGG 18

## RESULT 11

AAU25822

ID AAU25822 standard; peptide; 19 AA.

AC AAU25822;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #8.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KM haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KM bone marrow transplantation; haematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 OS tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;

XX Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Podduturi S;

XX yin Q;

XX WPI; 2001-564142/63.

XX Dielosure; Col 67-68; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods

XX of activating thrombopoietin receptors in cells comprise contacting the

XX cells with effective amounts of peptides and peptide mimetics attached to

XX hydrophilic polymers. The methods are used to treat thrombocytopenia such

XX as that due to chemotherapy, radiation therapy or bone-marrow

XX transplantation and to prevent thrombocytopenia in patients at risk. The

XX sequences are used to treat and prevent haematological disorders

XX including thrombocytopenia and platelet disorders. They are used in vitro

XX (TPO) and to develop other compounds that bind to and activate the TPO

XX receptor. The peptides can be used to detect TPO receptors on living

XX cells and fixed cells, in biological fluids, in tissue homogenates, and

XX in purified or natural biological materials. They may also be used for in

XX situ staining; fluorescence-activated cell sorting; Western blotting and

XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

XX be used for in vitro expansion of megakaryocytes and their committed

XX progenitors alone or in conjunction with additional cytokines

Best Local Similarity 100.0%; Pred. No. 4.8e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCADGPTLREWISFCGG 18  
 |||||  
 Db 1 GGCADGPTLREWISFCGG 18

## RESULT 12

AAW09466

ID AAW09466 standard; protein; 14 AA.

AC AAW09466;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound cyclic peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;

XX bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX Key Location/Qualifiers

XX Disulfide-bond 1. 14

XX Modified-site /note= "In acetyl form"

XX Modified-site 14

XX /note= "In amide form"

XX W09640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

XX 07-JUN-1995; 95US-00473604.

XX 07-JUN-1995; 95US-00476168.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00484090.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

XX mimetic(s) - useful in treatment of haematological disorders, esp.

XX thrombocytopenia resulting from chemotherapy, etc.

XX Claim 30; Page 91; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)

XX receptor (TR). The compound can be used for treating patients suffering

XX from haematological disorders and thrombocytopenia resulting from

XX chemotherapy, radiation therapy or bone marrow transfusions. The peptide

XX may also be used to maintain the proliferation and growth of TPO-

XX dependent cell lines and for use in biological research, for detecting

XX TPO receptors on living cells

XX Sequence 14 AA;

XX Query Match 76.0%; Score 85; DB 2; Length 14;

XX Best Local Similarity 100.0%; Pred. No. 1.3e-05;

XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Qy 3 CADGPTLREWISFC 16

XX Db 1 CADGPTLREWISFC 14

CC	XX	AAW09462	standard; protein; 14 AA.
CC	XX	AAW09462;	
CC	DT	10-SEP-1997	(first entry)
CC	XX	Thrombopoietin receptor binding compound peptide.	
CC	XX	Haematology; thrombocytopenia; TPO; TR; proliferation;	
CC	KW	bone marrow transfusion; chemotherapy; radiation therapy.	
CC	OS	Synthetic.	
CC	XX		
CC	XX	Key	Location/Qualifiers
CC	FT	Misc-difference	1..14
CC	FT		/note= "Preferably linkages are selected from: -
CC	FT		CH2OC(O)NR-; phosphate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
CC	FT		; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
CC	FT		lower alkyl"
CC	FT	Modified-site	1
CC	FT		/note= "Preferably N-terminus is selected from: -NR1; -
CC	FT		NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NR; succinimide;
CC	FT		benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3
CC	FT		substitutions on the phenyl ring selected from lower
CC	FT		alkyl, lower alkoxy, chloro, bromo; where R and R1 are
CC	FT		independently selected from hydrogen and lower alkyl"
CC	FT	Modified-site	14
CC	FT		/note= "Preferably C-terminus is -C(O)R2 where R2 is
CC	FT		selected from hydroxy, lower alkoxy, and -NR3R4, where R3
CC	FT		and R4 are independently selected from hydrogen and lower
CC	FT		alkyl, and where the nitrogen atom of the -NR3R4 group
CC	FT		can optionally be the amine group of the N-terminus of
CC	FT		the peptide forming a cyclic peptide"
CC	XX		
CC	XX	WO9640189-A1.	
CC	PD	19-DEC-1996.	
CC	XX		
CC	PF	05-JUN-1996;	96WC-US008998.
CC	XX		
CC	XX	07-JUN-1995;	95US-00472371.
CC	PR	07-JUN-1995;	95US-00473604.
CC	PR	07-JUN-1995;	95US-00476168.
CC	PR	07-JUN-1995;	95US-00478128.
CC	PR	07-JUN-1995;	95US-00484090.
CC	PR	07-JUN-1995;	95US-00485301.
CC	XX		
CC	PA	(GLAXO ) GLAXO GROUP LTD.	
CC	PI	Dower WJ, Barrett RW, Cwila SE, Duffin DJ, Gates CM, Johnson SS;	
CC	PI	Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;	
CC	XX		
CC	DR	WPI, 1997-051883/05.	
CC	XX		
CC	PT	Thrombopoietin receptor-binding/activating peptide(s) and peptide	
CC	PT	mimetic(s) - useful in treatment of haematological disorders, esp.	
CC	XX	thrombocytopenia resulting from chemotherapy, etc.	
CC	XX		
CC	PS	Claim 18; Page 89; 106pp; English.	
CC	XX		
CC	CC	The present sequence is a compound which binds to thrombopoietin (TPO)	
CC	CC	receptor (TR). It has a molecular weight of < 8000 Da, and a binding	
CC	CC	affinity to TR as expressed by an IC50 of no more than about 100 mM. The	
CC	CC	compound (especially if modified, see features table) can be used for	
CC	CC	treating patients suffering from haematological disorders and	
CC	CC	thrombocytopenia resulting from chemotherapy, radiation therapy or bone	
CC	CC	marrow transfusions. The peptide may also be used to maintain the	
CC	CC	proliferation and growth of TPO-dependent cell lines and for use in	
CC	CC	biological research, for detecting TPO receptors on living cells	

[illegible]



XX 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS,  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Disclosure; Page 26; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 14 AA;  
 Query Match 78.0%; Score 85; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14  
 RESULT 18  
 AAW33029  
 ID AAW33029 standard; peptide; 14 AA.  
 XX  
 AC AAW33029;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS,  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Claim 19; Page 89; 106pp; English.  
 XX  
 CC The present peptide binds the thrombopoietin receptor (TR), has a  
 CC molecular weight of less than 8000 Da and a TR binding affinity as  
 CC expressed by an IC50 of no more than about 100 microm. It can be used to  
 CC treat disorders which are susceptible to treatment with a thrombopoietin  
 CC agonist, preferably haematological disorders and thrombocytopenia  
 CC resulting from chemotherapy, radiation therapy or bone marrow  
 CC transfusions. It can also be used diagnostically, e.g. to investigate the  
 CC mechanism of thrombopoietin signal transduction and receptor activation,  
 CC or to maintain the proliferation and growth of thrombopoietin dependent  
 CC cell lines  
 XX  
 SQ Sequence 14 AA;  
 Query Match 78.0%; Score 85; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14  
 RESULT 19  
 AAW35401  
 ID AAW35401 standard; peptide; 14 AA.  
 XX  
 AC AAW35401;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FH Disulfide-bond 1..14  
 FT Modified-site 14  
 FT /note= "NH2-D-Cys"  
 XX  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS,  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 6; Page 63; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX  
 SQ Sequence 14 AA;

Query Match 78.0%; Score 85; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14

RESULT 20

AAW36647 ID AAW36647 standard; peptide; 14 AA.

XX AAW36647;

XX 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

KM signal transduction; receptor activation; cell culture.

XX Synthetic.

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 26; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX  
 SQ Sequence 14 AA;

Query Match 78.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CADGPTLRWISFC 16

Db |||||  
 1 CADGPTLRWISFC 14

RESULT 21

AAW35400 ID AAW35400 standard; peptide; 14 AA.

XX AAW35400;

XX 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

XX signal transduction; receptor activation; cell culture.

XX Synthetic.

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 6; Page 63; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX  
 SQ Sequence 14 AA;

Query Match 78.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CADGPTLRWISFC 16

Db 1 CADGPTLRWISFC 14

RESULT 22  
 AAW33032 ID AAW33032 standard; peptide; 14 AA.

```

XX AAW3032;
AC
XX
XX 11-MAR-1998 (first entry)
DT
XX
XX Thrombopoietin receptor binding peptide.
DE
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KM radiation therapy; bone marrow transfusion; diagnosis;
KM signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Disulfide-bond 1. .14
FT Modified-site 1
FT Modified-site /note= "acylated"
FT Modified-site /note= "amidated"
FT
XX
XX MO9640750-A1.
XX
XX 19-DEC-1996.
PD
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS,
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 30; Page 91; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transplants. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
XX Sequence 14 AA;
SQ
XX
XX Query Match 78.0%; Score 85; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CADGPTLRWISFC 16
DB 1 CADGPTLRWISFC 14

```

```

DE TPO-mimetic peptide sequence SEQ ID NO:70.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiaesthetic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTR44; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX thrombolysis; pharmaceutical.
XX
XX Synthetic.
OS
XX
XX MO200024782-A2.
XX
XX 04-MAY-2000.
PD
XX
XX 25-OCT-1999; 99WO-US025044.
XX
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.
XX
XX (AMGEN-) AMGEN INC.
XX
XX Peige U, Liu C, Cheatham J, Boone TC;
PI WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 19; Page 218; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiaesthetic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombolysis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions,
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
XX Sequence 14 AA;
SQ
XX
XX Query Match 78.0%; Score 85; DB 3; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CADGPTLRWISFC 16
DB 1 CADGPTLRWISFC 14

```

```

RESULT 24
AAU25826
ID AAU25826 standard; peptide; 14 AA.
XX
XX AAU25826;
XX
XX 17-DEC-2001 (first entry)
DT
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #12.
DE
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
XX

```



XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iacti gene.  
 OS Homo sapiens.  
 XX US6251864-B1.  
 PN 26-JUN-2001.  
 XX 01-MAR-2000; 2000US-00516704.  
 PF 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX (GLAXO ) GLAXO GROUP LTD.  
 PA Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RM, Depirince RB, Poddaturi S;  
 PI Yan Q;  
 XX WPI; 2001-564142/63.  
 DR Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 XX with peptides and peptide mimetics attached to hydrophilic polymers.  
 PT Disclosure; Col 20; 128pp; English.  
 PS Sequences AM25815-AM26049 represent peptides and peptide mimetics that  
 XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX Sequence 14 AA;  
 SQ Query Match 78.0%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 RESULT 27  
 ABB72900  
 ID ABB72900 standard; peptide; 14 AA.  
 XX ABB72900;  
 AC ABB72900;  
 XX 05-APR-2002 (first entry)  
 DT

DE TPO mimetic peptide SEQ ID NO:70.  
 XX Modified peptide, mimetic; Fe domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EGF; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumor; immunosuppressive;  
 KW cytostatic; antineumatic; antitachytic; antidiabetic; ophthalmological;  
 KW antianemic; anorectic; antifertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200183525-A2.  
 PN 08-NOV-2001.  
 PD 02-MAY-2001; 2001WO-US014310.  
 PF 03-MAY-2000; 2000US-00563286.  
 PR (AMGE-) AMGEN INC.  
 XX Feige U, Liu C, Cheatham JC, Boone TC, Gudae JM;  
 PI WPI; 2002-130313/17.  
 DR Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 XX diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX Claim 39; Page 44; 176pp; English.  
 PS The present invention describes a vehicle-peptide molecule (I) or its  
 XX multimers. (I) can have antiinflammatory, antitumor, immunosuppressive,  
 CC cytostatic, antineumatic, antitachytic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction or their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX Sequence 14 AA;  
 SQ Query Match 78.0%; Score 85; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 RESULT 28



ADJ73051 standard; peptide: 14 AA.  
 ID ADJ73051;  
 XX  
 AC ADJ73051;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE TPO mimetic peptide sequence SeqID 505.  
 XX  
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
 KW TPO.  
 XX  
 OS Synthetic.  
 OS  
 PN WO2003084477-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 24-MAR-2003; 2003WO-US009139.  
 XX  
 PR 29-MAR-2002; 2002US-0368791P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Scallion BJ, Ghraryeb J;  
 XX  
 DR WPI; 2003-804237/75.  
 XX  
 PT New CDR mimetibody comprising a portion of a heavy or light chain  
 PT variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PT cardiovascular or neurologic disease.  
 XX  
 PS Disclosure; SEQ ID NO 505; 97pp; English.  
 XX  
 CC This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LBR). The present invention describes human  
 CC mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neurologic  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 CC  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 78.0%; Score 85; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLREWISFC 16  
 DB 1 CADGPTLREWISFC 14  
 XX  
 RESULT 29  
 ADJ52686 standard; peptide: 14 AA.  
 ID ADJ52686;  
 XX  
 AC ADJ52686;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX

CH1 deleted mimetibody-related peptide SeqID505.  
 DE  
 XX  
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 KW hypotensive; neuroprotective; nootropic; antibacterial; virocidic;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX  
 OS Unidentified.  
 OS  
 OS Synthetic.  
 OS  
 PN WO2004002417-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 27-JUN-2003; 2003WO-US020347.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Ghraryeb J, Scallion BJ, Neespor TC;  
 XX  
 PI Kutoloshki KA;  
 XX  
 DR WPI; 2004-082870/08.  
 XX  
 PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 XX  
 PS Claim 2; SEQ ID NO 505; 129pp; English.  
 XX  
 CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virocidic or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 CC  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 78.0%; Score 85; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLREWISFC 16  
 DB 1 CADGPTLREWISFC 14  
 XX  
 RESULT 30  
 ADJ51647 standard; peptide: 14 AA.  
 ID ADJ51647;  
 XX  
 AC ADJ51647;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CH1 deleted mimetibody-related peptide SeqID505.  
 XX  
 KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstetric disorder; haematologic disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 KW Unidentified.  
 OS Synthetic.  
 PN WO2004002424-A2.  
 XX  
 XX 08-JAN-2004.  
 PD  
 XX 30-JUN-2003; 2003WO-US020495.  
 PF  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR 13-SEP-2002; 2002US-0412144P.  
 XX  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 PI Heaver GA, Knight DM, Ghayeb J, Scallion BJ, Nespor TC;  
 PI Kutolooski KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 PT  
 XX  
 PS Claim 15, SEQ ID NO 505; 123pp; English.  
 XX  
 XX This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, allergic, infectious,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 78.0%; Score 85; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14

ID AAM09467 standard; protein; 13 AA.  
 XX  
 XX AAM09467;  
 AC  
 XX 10-SEP-1997 (first entry)  
 DT  
 XX Thrombopoietin receptor binding compound cyclic peptide.  
 DE  
 KW Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;  
 KW bone marrow transfusion; chemotherapy; radiation therapy.  
 XX  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note="The Ala is linked with the modified Cys at  
 FT position 13"  
 FT  
 FT Modified-site 14  
 FT /label= OTHER  
 FT /note="S-carboxymethyl-cysteine alpha-carboxamide;  
 FT forming a linkage onto the Ala at position one with the  
 FT delta C of this residue"  
 XX  
 XX WO9640189-A1.  
 PN  
 XX 19-DEC-1996.  
 PD  
 XX 05-JUN-1996; 96WO-US008998.  
 PF  
 XX 07-JUN-1995; 95US-00472371.  
 PR 07-JUN-1995; 95US-00473604.  
 PR 07-JUN-1995; 95US-00476168.  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00484090.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 PI  
 XX WPI; 1997-051883/05.  
 DR  
 XX Thrombopoietin receptor-binding/activating peptide(s) and peptide  
 PT mimetic(s) - useful in treatment of haematological disorders, esp.  
 PT thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 XX Claim 30; Page 91; 106pp; English.  
 PS  
 XX The present sequence is a compound which binds to thrombopoietin (TPO)  
 CC receptor (TR). The compound can be used for treating patients suffering  
 CC from haematological disorders and thrombocytopenia resulting from  
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide  
 CC may also be used to maintain the proliferation and growth of TPO-  
 CC dependent cell lines and for use in biological research, for detecting  
 CC TPO receptors on living cells  
 CC  
 XX  
 SQ Sequence 13 AA;  
 Query Match 69.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 Db 1 ADGPTLRWISFC 13

RESULT 31  
 AAM09467

RESULT 32  
 AAM35399  
 ID AAM35399 standard; peptide; 13 AA.  
 XX  
 AC AAM35399;

```

XX 11-MAR-1998 (first entry)
DT Thrombopoietin receptor binding peptide.
DE Thrombopoietin receptor binding peptide.
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopaenia; chemotherapy;
KM radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX Synthetic.
XX Key Location/Qualifiers
FT Modified-site 1 /note= "COCH2-alanine linked via CH2 group to Cys13"
FT Modified-site 13 /note= "NH2-cytosine linked via sulphoxidised thiol group
FT to Ala1"
XX WO9640750-A1.
XX 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009623.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX (GLAXO ) GLAXO GROUP LTD.
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX Example 6; Page 63; 106pp; English.
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX Sequence 13 AA;
SQ
Query Match 69.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ADGPTLRWISFC 16
DB 1 ADGPTLRWISFC 13

```

```

KM radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX Synthetic.
XX Key Location/Qualifiers
FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical
FT peptide"
FT Modified-site 13 /note= "NH2-Phe"
XX WO9640750-A1.
XX 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009623.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX (GLAXO ) GLAXO GROUP LTD.
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX Example 9; Page 73; 106pp; English.
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX Sequence 13 AA;
SQ
Query Match 69.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CADGPTLRWISF 15
DB 1 CADGPTLRWISF 13

```

```

RESULT 33
AAW35417
ID AAW35417 standard; peptide; 13 AA.
XX AAW35417;
AC AAW35417;
XX 11-MAR-1998 (first entry)
DT Thrombopoietin receptor binding peptide.
DE Thrombopoietin receptor binding peptide.
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;

```

```

RESULT 34
AAW33033
ID AAW33033 standard; peptide; 13 AA.
XX AAW33033;
AC AAW33033;
XX 11-MAR-1998 (first entry)
DT Thrombopoietin receptor binding peptide.
DE Thrombopoietin receptor binding peptide.
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX Synthetic.
XX Key Location/Qualifiers
FT Modified-site 1

```

FT	Modified-site	/note= "COCH2-alanine linked via CH2 group to Cys13"
FT	13	
FT	/note= "NH2-cytosine linked via thiol group to Ala1"	
XX		
PN	WO9640750-A1.	
XX		
PD	19-DEC-1996.	
XX		
PF	07-JUN-1996;	96WO-US0009623.
XX		
PR	07-JUN-1995;	95US-00478128.
PR	07-JUN-1995;	95US-00485301.
XX		
PA	(GLAX ) GLAXO GROUP LTD.	
XX		
P1	Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;	
P1	Mattheakis LC, Schatz PJ, Magstrom CR, Wrighton NC;	
XX		
DR	WPI; 1997-052226/05.	
XX		
PT	Peptides and peptide mimetics which bind to and activate the	
PT	thrombopoietin receptor - useful in treatment of haematological	
PT	disorders, esp. thrombocytopenia resulting from chemotherapy, etc.	
XX		
PS	Claim 30; Page 91; 106pp; English.	
XX		
CC	The present peptide binds the thrombopoietin receptor (TR), has a	
CC	molecular weight of less than 8000 Da and a TR binding affinity as	
CC	expressed by an IC50 of no more than about 100 microm, it can be used to	
CC	treat disorders which are susceptible to treatment with a thrombopoietin	
CC	agonist, preferably haematological disorders and thrombocytopaenia	
CC	resulting from chemotherapy, radiation therapy or bone marrow	
CC	transfusions. It can also be used diagnostically, e.g. to investigate the	
CC	mechanism of thrombopoietin signal transduction and receptor activation,	
CC	or to maintain the proliferation and growth of thrombopoietin dependent	
CC	cell lines	
XX		
S0	Sequence 13 AA;	
Query Match	69.7%; Score 76; DB 2; Length 13;	
Best Local Similarity	100.0%; Pred. No. 0.00025;	
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	4 ADGPTLRMISFC 16	
DB	1 ADGPTLRMISFC 13	
RESULT 35		
AAW35413		
ID	AAW35413 standard; peptide; 13 AA.	
XX		
AC	AAW35413;	
XX		
DT	11-MAR-1998 (first entry)	
XX		
DE	Thrombopoietin receptor binding peptide.	
XX		
KW	Thrombopoietin receptor; binding peptide; treatment; agonist;	
KW	haematological disorder; thrombocytopaenia; chemotherapy;	
KW	radiation therapy; bone marrow transfusion; diagnosis;	
KW	signal transduction; receptor activation; cell culture.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	1
FT	/note= "Br-Ala"	
FT	Modified-site	13
FT	/note= "NH2-Cys"	
XX		
PN	WO9640750-A1.	
XX		

[illegible]

XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX Example 6; Page 64; 106pp; English.  
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 13 AA;  
 Query Match 69.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13  
 Db  
 RESULT 37  
 ID AAW35422  
 AAW35422 standard; peptide; 13 AA.  
 AC AAW35422;  
 XX 11-MAR-1998 (first entry)  
 DT  
 XX Thrombopoietin receptor binding peptide.  
 DE Thrombopoietin receptor; binding peptide; treatment; agonist;  
 XX haematological disorder; thrombocytopenia; chemotherapy;  
 KM radiation therapy; bone marrow transfusion; diagnosis;  
 KM signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "optionally acylated"  
 FT Cross-links 13  
 FT /note= "linked via disulfide bond to Cys13 of identical  
 peptide"  
 FT  
 FT  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX Example 9; Page 74; 106pp; English.  
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 13 AA;  
 Query Match 69.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13  
 Db  
 RESULT 38  
 ID AAW35397  
 AAW35397 standard; peptide; 13 AA.  
 AC AAW35397;  
 XX 11-MAR-1998 (first entry)  
 DT  
 XX Thrombopoietin receptor binding peptide.  
 DE Thrombopoietin receptor; binding peptide; treatment; agonist;  
 XX haematological disorder; thrombocytopenia; chemotherapy;  
 KM radiation therapy; bone marrow transfusion; diagnosis;  
 KM signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "COCH2-alanine linked via CH2 group to Cys13"  
 FT Modified-site 13  
 FT /note= "NH2-cytosine linked via thiol group to Ala1"  
 FT  
 FT  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX Example 6; Page 63; 106pp; English.  
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transplants. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

SQ Sequence 13 AA;

Query Match 69.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13

RESULT 39

AAU25997  
 ID AAU25997 standard; peptide; 13 AA.

AC AAU25997;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #183.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 XX bone marrow transplantation; haematological disorder; platelet disorder;  
 XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iaci gene.

OS Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,

PI Balaubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S,

PI Yin Q;

DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 143-144; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

SQ Sequence 13 AA;

Query Match 69.7%; Score 76; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CADGPTLRWISF 15  
 |||||  
 1 CADGPTLRWISF 13

RESULT 40

AAU25984  
 ID AAU25984 standard; peptide; 13 AA.

AC AAU25984;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #170.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 XX bone marrow transplantation; haematological disorder; platelet disorder;  
 XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iaci gene.

OS Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,

PI Balaubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S,

PI Yin Q;

DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 137; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

Sequence 13 AA;

Query Match 69.7%; Score 76; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ADGPTLRWISFC 16  
| | | | | | | | | | | | |  
Db 1 ADGPTLRWISFC 13

#### RESULT 41

AAW35398 standard; peptide; 14 AA.

AAW35398;  
11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;  
haematological disorder; thrombocytopenia; chemotherapy;  
radiation therapy; bone marrow transfusion; diagnosis;  
signal transduction; receptor activation; cell culture.

Synthetic.

Key Location/Qualifiers  
FT Disulfide-bond 1..14  
FT Modified-site /note= "Homocysteine"  
FT Modified-site 14  
FT Modified-site /note= "NH2-Cys"

W09640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAXO) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the thrombopoietin receptor - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Example 6; Page 63; 106pp; English.

The present peptide, which binds the thrombopoietin receptor (TR), can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

Sequence 14 AA;

Query Match 69.7%; Score 76; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ADGPTLRWISFC 16  
| | | | | | | | | | | | |  
Db 2 ADGPTLRWISFC 14

#### RESULT 42

AAW35396 standard; peptide; 14 AA.

AAW35396;

11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;  
haematological disorder; thrombocytopenia; chemotherapy;  
radiation therapy; bone marrow transfusion; diagnosis;  
signal transduction; receptor activation; cell culture.

Synthetic.

Key Location/Qualifiers  
FT Disulfide-bond 1..14  
FT Modified-site /note= "Penicillamine"  
FT Modified-site 14  
FT Modified-site /note= "NH2-Cys"

W09640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAXO) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the thrombopoietin receptor - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Example 6; Page 63; 106pp; English.

The present peptide, which binds the thrombopoietin receptor (TR), can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

Sequence 14 AA;

Query Match 69.7%; Score 76; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ADGPTLREWISFC 16  
 |||||  
 DB 2 ADGPTLREWISFC 14

## RESULT 43

AAW35402  
 ID AAW35402 standard; peptide; 14 AA.

XX AAW35402;

AC 11-MAR-1998 (first entry)

DT Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX Synthetic.

OS Key Location/Qualifiers

FT Disulfide-bond 1. .14

FT Modified-site /note= "D-form residue, Penicillamine"

FT Modified-site 14

FT /note= "NH2-D-Cys"

PN WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstroom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 6; Page 64; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX Sequence 14 AA;

XX SQ

Query Match 69.7%; Score 76; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00077;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ADGPTLREWISFC 16  
 |||||  
 DB 2 ADGPTLREWISFC 14

ID AAU25987 standard; peptide; 14 AA.

XX AAU25987;

AC 18-DEC-2001 (first entry)

DT Human thrombopoietin receptor (TPO-R) activator peptide #173.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

KW bone marrow transplantation; haematological disorder; platelet disorder;

KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;

KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;

KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAXO ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstroom CR, Hendren RW, Depirince RB, Poddaturi S;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 139; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

XX progenitors alone or in conjunction with additional cytokines

XX Sequence 14 AA;

XX SQ

Query Match 69.7%; Score 76; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00027;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CADGPTLREWISF 15  
 |||||  
 DB 1 CADGPTLREWISF 13

RESULT 44  
 AAU25987

RESULT 45



AAU25983  
ID AAU25983 standard; peptide; 14 AA.

AC AAU25983;

DT 18-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #169.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
KM haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
KM bone marrow transplantation; haematological disorder; platelet disorder;  
KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iaci gene.

OS Homo sapiens.

PN US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO ) GLAXO GROUP LTD.

PI Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz RJ,

PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Podduturi S;

PI Yin Q;

DR WPI; 2001-564142/63.

XX

PT Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 135-137, 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

CC progenitors alone or in conjunction with additional cytokines

XX

SQ Sequence 14 AA;

Query Match 69.7%; Score 76; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00027; Mismatches 0; Gaps 0;

Matches 13; Conservative 0; Indels 0;

QY 4 ADGPTLRWISFC 16

DB 2 ADGPTLRWISFC 14

Search completed: September 1, 2005, 16:12:08  
Job time : 84.7482 secs

This Page Blank (usps)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:57:33 ; Search time 13.7266 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-6

Perfect score: 109  
Sequence: 1 GGCAGDPTLRMTSFCG 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

1: PIR.79.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	45.0	245	2 T47701	translatation initia
2	47	43.1	475	2 T33943	hypothetical prote
3	45	41.3	108	2 T49731	hypothetical prote
4	45	41.3	180	2 T4944	hypothetical prote
5	45	41.3	421	2 T22969	hypothetical prote
6	45	41.3	499	2 S51089	ammonium transport
7	44	40.4	346	2 T19008	hypothetical prote
8	44	40.4	346	2 A58583	testosterone-resis
9	44	40.4	371	2 D75266	cell division prot
10	44	40.4	490	2 T09084	phosphatidylinosit
11	44	40.4	526	2 A86440	58.5K hypothetical
12	44	40.4	974	2 S34189	starch phosphoryla
13	44	40.4	1022	1 S00503	Na+/K+-exchanging
14	44	40.4	1023	2 A24414	Na+/K+-exchanging
15	43.5	39.9	376	2 T39685	conserved hypothet
16	43.5	39.9	1499	2 A89813	glutamate synthase
17	43	39.4	113	2 D72595	hypothetical prote
18	43	39.4	115	2 T15386	hypothetical prote
19	43	39.4	230	2 A18685	lipote biosynthes
20	43	39.4	233	2 A82768	lipote biosynthes
21	43	39.4	246	2 T19988	hypothetical prote
22	43	39.4	247	2 T01012	probable translati
23	43	39.4	268	2 D97548	lipote-protein li
24	43	39.4	276	2 A38654	maet cell proteina
25	43	39.4	953	2 S54478	probable membrane
26	43	39.4	1010	2 B37227	Na+/K+-exchanging
27	43	39.4	1013	1 S00801	Na+/K+-exchanging
28	43	39.4	1013	2 C24639	Na+/K+-exchanging
29	43	39.4	1017	2 A37227	Na+/K+-exchanging

30	43	39.4	1020	2 A34474	Na+/K+-exchanging
31	43	39.4	1020	2 B24639	Na+/K+-exchanging
32	43	39.4	1021	1 PMSHNA	Na+/K+-exchanging
33	43	39.4	1021	1 S04630	Na+/K+-exchanging
34	43	39.4	1021	1 A28139	Na+/K+-exchanging
35	43	39.4	1021	2 B24862	Na+/K+-exchanging
36	43	39.4	1022	2 S49127	Na+/K+-exchanging
37	43	39.4	1023	1 A24639	Na+/K+-exchanging
38	43	39.4	1023	1 S24650	Na+/K+-exchanging
39	43	39.4	1025	2 A60444	Na+/K+-exchanging
40	43	39.4	1027	1 PMCNM	Na+/K+-exchanging
41	43	39.4	1038	1 S03632	hypothetical prote
42	42.5	39.0	353	2 T32638	hypothetical prote
43	42.5	39.0	1004	2 JH0470	Na+/K+-exchanging
44	42.5	39.0	1302	2 T00038	hypothetical prote
45	42	38.5	141	2 AH2829	conserved hypothet
46	42	38.5	141	2 F97607	hypothetical prote
47	42	38.5	192	1 A24902	erythropoietin pre
48	42	38.5	192	1 S28148	erythropoietin pre
49	42	38.5	312	2 F86876	hypothetical prote
50	42	38.5	440	2 F81555	glutamate-1-semial
51	42	38.5	440	2 B86508	glutamate-1-semial
52	42	38.5	440	2 G72114	glutamate-1-semial
53	42	38.5	473	2 T31717	hypothetical prote
54	42	38.5	522	2 D69226	hypothetical prote
55	42	38.5	522	2 S62941	probable membrane
56	42	38.5	725	2 A11544	conserved hypothet
57	42	38.5	842	2 T12091	starch phosphoryla
58	41.5	38.1	108	2 G82991	chlorodoxin PAS240
59	41	37.6	132	1 G69256	conserved hypothet
60	41	37.6	189	2 S07755	hypothetical prote
61	41	37.6	245	2 JC7273	inducible mast cel
62	41	37.6	273	2 H70849	hypothetical prote
63	41	37.6	274	2 A45754	tryptase (EC 3.4.2
64	41	37.6	275	2 C35863	tryptase (EC 3.4.2
65	41	37.6	298	2 T23362	hypothetical prote
66	41	37.6	410	1 DBPSXA	3-methyl-2-oxobuta
67	41	37.6	410	2 C83365	2-oxoisovalerate d
68	41	37.6	473	2 E84853	hypothetical prote
69	41	37.6	494	2 H82489	conserved hypothet
70	41	37.6	576	2 C88950	protein R0985.11 f
71	41	37.6	593	2 S45281	coagulation factor
72	41	37.6	618	2 T48193	hypothetical prote
73	41	37.6	929	2 S75098	hypothetical prote
74	41	37.6	955	2 T10947	starch phosphoryla
75	41	37.6	966	1 PHP0AG	starch phosphoryla
76	41	37.6	971	2 T09210	starch phosphoryla
77	41	37.6	1000	2 S47243	starch phosphoryla
78	41	37.6	1313	2 B96509	protein F27F5.11 f
79	41	37.6	1522	2 C96578	hypothetical prote
80	41	37.6	1616	2 T17884	S-layer protein -
81	40.5	37.2	1363	2 T43320	insulin-like growt
82	40	36.7	98	2 A70301	ribosomal protein
83	40	36.7	152	2 S21826	T-cell receptor be
84	40	36.7	155	2 S23629	hypothetical prote
85	40	36.7	157	2 B83066	hypothetical prote
86	40	36.7	169	1 ICMS2	interleukin-2 prec
87	40	36.7	169	2 S37289	interleukin-2 prec
88	40	36.7	169	2 B95908	hypothetical prote
89	40	36.7	188	2 T33623	cytochrome-c oxida
90	40	36.7	206	2 E45315	hypothetical prote
91	40	36.7	206	2 T22345	pol polyprotein -
92	40	36.7	217	2 S46354	lipote-protein li
93	40	36.7	226	2 G87518	hypothetical prote
94	40	36.7	241	2 B83447	hypothetical prote
95	40	36.7	252	2 B97072	probable hydrolase
96	40	36.7	357	2 T37154	hypothetical prote
97	40	36.7	360	2 S25561	transcription fact
98	40	36.7	361	2 F91207	hypothetical prote
99	40	36.7	361	2 H86053	hypothetical prote
100	40	36.7	361	2 C65171	hypothetical 41.0

## ALIGNMENTS

## RESULT 1

T47701 translation initiation factor eif-6-like protein [imported] - Arabidopsis thaliana

N/Alternate names: protein F116.30

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C/Accession: T47701

R/Sensu: V.; Wurmbach, E.; Drzonek, H.; Anseorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.

submitted to the Protein Sequence Database, March 2000

A/Reference number: Z24473

A/Accession: T47701

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-245 &lt;BEN&gt;

A/Cross-references: UNIPROT:Q9M060; EMBL:AL161667

A/Experimental source: cultivar Columbia; BAC clone F116

C/Genetics:

A/Map position: 3

A/Introns: 4/1; 36/2; 65/1; 80/1; 123/3; 160/3

A/Note: F116.30

C/Superfamily: conserved hypothetical protein YP0016c

## Query Match

45.0%; Score 49; DB 2; Length 245;

Best Local Similarity 57.1%; Pred. No. 8.2;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 ADGPTLRWISFCG 17

DB 194 AAGMTVNDWTSFCG 207

## RESULT 2

T33943 hypothetical protein C01B4.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T33943

R/Smith, A.; Wamsley, P.; Fromick, W.

submitted to the EMBL Data Library, February 1999

A/Description: The sequence of C. elegans cosmid C01B4.

A/Reference number: Z21443

A/Accession: T33943

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-475 &lt;SMI&gt;

A/Cross-references: UNIPROT:Q9UAT5; EMBL:AF125952; PIDD:AA014699.1; GSPDB:GN00023; CESP:

A/Experimental source: strain Bristol N2; clone C01B4

C/Genetics:

A/Map position: 5

A/Introns: 45/2; 80/1; 118/2; 189/3; 239/2; 340/3; 433/3

## Query Match

43.1%; Score 47; DB 2; Length 475;

Best Local Similarity 50.0%; Pred. No. 30;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCG 18

DB 268 CTDRCVLSAMVSLDGG 283

## RESULT 3

T49731 hypothetical protein B24B19.30 [imported] - Neurospora crassa

C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000

C/Accession: T49731

R/Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25022

A/Accession: T49731

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-108 &lt;SCH&gt;

A/Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.30

A/Experimental source: BAC clone B24B19; strain OR74A

C/Genetics:

A/Map position: 6

C/Superfamily: Neurospora crassa hypothetical protein B24B19.30

## Query Match

41.3%; Score 45; DB 2; Length 108;

Best Local Similarity 50.0%; Pred. No. 15;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCG 16

DB 70 CCGQPTLRWMLSMC 83

## RESULT 4

T44944 hypothetical protein 5 [imported] - Natronobacterium pharaonis

C/Species: Natronobacterium pharaonis

C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C/Accession: T44944

R/Mattar, S.; Engelhard, W.

Eur. J. Biochem. 250, 332-341, 1997

A/Title: Cytochrome b3 from Natronobacterium pharaonis: An archaeal four-subunit cyto

A/Reference number: Z22876; PMID:9808958; PMID:9428682

A/Accession: T44944

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-180 &lt;MAT&gt;

A/Cross-references: UNIPROT:Q07291; EMBL:Y10500; PIDD:CAA71527.1

A/Experimental source: strain SP1/28

C/Genetics:

A/Map position: 5

C/Superfamily: conserved hypothetical protein AF1745

## Query Match

41.3%; Score 45; DB 2; Length 180;

Best Local Similarity 77.8%; Pred. No. 24;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 LREWISFCG 17

DB 116 LLEWLSFCG 124

## RESULT 5

T22969 hypothetical protein F59A1.13 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T22969

R/Mortimore, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19644

A/Accession: T22969

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-421 &lt;WIL&gt;

A/Cross-references: UNIPROT:Q9XUV7; EMBL:Z81557; PIDD:CA04538.1; GSPDB:GN00023; CESP:

A/Experimental source: clone F59A1

C/Genetics:

A/Map position: 13

A/Introns: 27/1; 116/1; 245/3; 286/3; 340/3; 381/3

## Query Match

41.3%; Score 45; DB 2; Length 421;

Best Local Similarity 50.0%; Pred. No. 53;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 3 CADGPTLRWISFCG 18  
 Db 214 CTDGTVLWGLSVFVG 229

## RESULT 6

ammonium transport protein MEP2 - yeast (Saccharomyces cerevisiae)  
 N/Alternate names: NH3 permease; protein JTA499; protein NI207; protein NI820; protein X  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 10-May-1995 #sequence revision 19-Oct-1995 #text\_change 09-Jul-2004  
 C/Accession: S51089, S55142, S59247, S63087  
 R/Martin, A.M.; Andre, B.  
 Submitted to the EMBL Data Library, December 1994  
 A/Reference number: S51089  
 A/Accession: S51089  
 A/Molecule type: DNA  
 A/Residues: 1-499 <MAB>  
 A/Cross-references: UNIPROT:P41948, EMBL:X83608; NID:G619513; PIDN:CAA58587.1; PID:G6195  
 R/Mallet, L.; Buserreau, F.; Jacquet, M.  
 Submitted to the EMBL Data Library, November 1994  
 A/Description: A 43.5 kb fragment of the chromosome XIV.  
 A/Reference number: S55136  
 A/Accession: S55142  
 A/Molecule type: DNA  
 A/Residues: 1-499 <MAL>  
 A/Cross-references: EMBL:Z46843; NID:G861113; PIDN:CAA66884.1; PID:G854496  
 R/Mallet, L.; Buserreau, F.; Jacquet, M.  
 Yeast 11, 1195-1209, 1995  
 A/Title: A 43.5 kb segment of yeast chromosome XIV, which contains MPA2, MEP2, CAP/SRV2.  
 A/Reference number: S59241; MUID:96109932; PMID:8619318  
 A/Accession: S59247  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-499 <MAW>  
 A/Cross-references: EMBL:Z46843; NID:G861113; PIDN:CAA66884.1; PID:G854496  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994  
 R/Mallet, L.; Buserreau, F.; Jacquet, M.  
 submitted to the Protein Sequence Database, April 1996  
 A/Reference number: S63069  
 A/Accession: S63087  
 A/Molecule type: DNA  
 A/Residues: 1-499 <MAF>  
 A/Cross-references: EMBL:Z71418; NID:G1302090; PIDN:CAA96025.1; PID:G1302091; MIBS:YNL14  
 A/Experimental source: strain S288C  
 C/Genetics:  
 A/Gene: SGD:MEP2  
 A/Cross-references: SGD:S0005086; MIBS:YNL142w  
 A/Map position: 14L  
 C/Function:  
 A/Description: ammonium transport  
 C/Superfamily: ammonium transport protein  
 C/Keywords: ammonium transport; transmembrane protein  
 F/35-51/Domain: transmembrane #status predicted <TM1>  
 F/62-78/Domain: transmembrane #status predicted <TM2>  
 F/123-139/Domain: transmembrane #status predicted <TM3>  
 F/154-170/Domain: transmembrane #status predicted <TM4>  
 F/228-244/Domain: transmembrane #status predicted <TM5>  
 F/288-304/Domain: transmembrane #status predicted <TM6>  
 F/306-332/Domain: transmembrane #status predicted <TM7>  
 F/397-413/Domain: transmembrane #status predicted <TM8>

Query Match 41.3%; Score 45; DB 2; Length 499;  
 Best Local Similarity 38.5%; Pred. No. 61;  
 Matches 10; Conservative 1; Mismatches 7; Indels 8; Gaps 1;

Oy 1 GGCAAGPTLRWISF-----CGG 18  
 Db 247 GGSAGNATIRAWYSIMSTNLAAACGG 272

RESULT 7  
 T19008

hypothetical protein C06C6.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T19008  
 R/McMurray, A.  
 Submitted to the EMBL Data Library, March 1997  
 A/Reference number: Z19059  
 A/Accession: T19008  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-346 <WIL>  
 A/Cross-references: UNIPROT:O62030; EMBL:Z93374; PIDN:CA807554.1; GSPDB:GN00023; CESP:C  
 A/Experimental source: clone C06C6  
 C/Genetics:  
 A/Gene: CESP:C06C6.2  
 A/Map position: 5  
 A/Intons: 109/1; 135/2; 160/2; 310/1  
 C/Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 40.4%; Score 44; DB 2; Length 346;  
 Best Local Similarity 56.2%; Pred. No. 61;  
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 2 GCADGPTLRWISFCG 17  
 Db 183 GLADGTTTMDSFIC 198

## RESULT 8

A58583  
 testosterone-resistant immunity-associated protein IAP38 - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 25-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
 C/Accession: A58583  
 R/Kruceken, J.; Schmitt-Wrede, H.P.; Markmann-Mullisch, U.; Wunderlich, F.  
 Biochem. Biophys. Res. Commun. 230, 167-170, 1997  
 A/Title: Novel gene expressed in spleen cells mediating acquired testosterone-resistant  
 A/Reference number: A58583; MUID:97146595; PMID:9020038  
 A/Accession: A58583  
 A/Molecule type: mRNA  
 A/Residues: 1-346 <KRU>  
 A/Cross-references: UNIPROT:P70224; GB:Y08026; NID:G1550784; PIDN:CAA69283.1; PID:G1550  
 A/Experimental source: spleen cell  
 C/Comment: This protein is a plasma membrane protein with two membrane-spanning domains  
 chabaudi malaria.  
 C/Genetics:  
 A/Gene: iap38  
 F/148-167/Domain: transmembrane #status predicted <TM1>  
 F/320-335/Domain: transmembrane #status predicted <TM2>

Query Match 40.4%; Score 44; DB 2; Length 346;  
 Best Local Similarity 43.8%; Pred. No. 61;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Oy 3 CADGPTLRWISFCG 18  
 Db 213 CTDNRALRDVAVACGG 228

## RESULT 9

D75266  
 cell division protein, FtsW/RodA/SpoVE family - Deinococcus radiodurans (strain R1)  
 C/Species: Deinococcus radiodurans  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: D75266  
 R/White, O.; Eissen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; M  
 Science 286, 1571-1577, 1999  
 A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A/Reference number: A75250; MUID:20036896; PMID:10567266  
 A/Accession: D75266  
 A/Status: preliminary

A;Molecule type: DNA  
 A;Residues: 1-371 <WHI>  
 A;Cross-references: UNIPROT:Q9RRJ3; GB:AE002079; GB:AE000513; NID:96460315; PIDN:AAF1203  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR2497  
 A;Map position: 1  
 C;Superfamily: rod shape-determining protein

Query Match 40.4%; Score 44; DB 2; Length 371;  
 Best Local Similarity 43.8%; Pred. No. 66;  
 Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCADGPTLRWISFCG 17  
 DB 77 GSGDSEGVRRWLSIAG 92

## RESULT 10

T09084  
 phosphatidylinositol 3-kinase - Chlamydomonas reinhardtii (fragment)  
 C;Species: Chlamydomonas reinhardtii  
 C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C;Accession: T09084  
 R;Molendijk, A.J.; Irvine, R.P.  
 Plant Mol. Biol. 37, 53-66, 1998  
 A;Title: Inositolide signalling in Chlamydomonas: Characterization of a phosphatidylinositol  
 A;Reference number: Z16411; PMID:98281574; PMID:9620264  
 A;Accession: T09084  
 A;Status: preliminary; translated from GB/EMBL/DDB  
 A;Molecule type: DNA  
 A;Residues: 1-490 <MOL>  
 A;Cross-references: UNIPROT:O04270; EMBL:U97663; NID:92109290; PIDN:AAC50018.1; PID:9210  
 A;Experimental source: strain cw-15  
 C;Genetics:  
 A;Introns: 265/3; 331/3; 370/3; 455/1; 481/3

Query Match 40.4%; Score 44; DB 2; Length 490;  
 Best Local Similarity 50.0%; Pred. No. 85;  
 Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

QY 1 GGCA--DGPTLR--EWISFC 16  
 DB 244 GSSSGDSEGVRRWLSIAG 263

## RESULT 11

A86440  
 58.5K hypothetical protein - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C;Accession: A86440  
 R;Thellogris, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 anson, N.P.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; PMID:21016719; PMID:11130712  
 A;Accession: A86440

A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-526 <STO>  
 A;Cross-references: UNIPROT:Q9C868; GB:AE005172; NID:911054679; PIDN:AA627899.1; GSPDB:Q  
 C;Genetics:  
 A;Map position: 1

Query Match 40.4%; Score 44; DB 2; Length 526;  
 Best Local Similarity 44.4%; Pred. No. 91;

Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 GCADGPT--LRWISFC 16  
 DB 395 GGRVGGPSPDLINQWIEFC 412

## RESULT 12

S34189  
 starch phosphorylase (EC 2.4.1.1) L - potato  
 C;Species: Solanum tuberosum (potato)  
 C;Date: 03-Mar-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C;Accession: S53489; S34189  
 R;Somewald, U.; Baener, A.; Greve, B.; Stemp, M.  
 Plant Mol. Biol. 27, 567-576, 1995  
 A;Title: A second L-type isozyme of potato glycan phosphorylase: cloning, antisense inh  
 A;Reference number: S53489; PMID:95201249; PMID:7894019  
 A;Accession: S53489  
 A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA  
 A;Residues: 1-974 <STO>  
 A;Cross-references: UNIPROT:P53535; EMBL:X73684; NID:9313348; PIDN:CAA52036.1; PID:93133  
 C;Superfamily: glucan phosphorylase  
 C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphat  
 F;820/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 40.4%; Score 44; DB 2; Length 974;  
 Best Local Similarity 58.3%; Pred. No. 1,6e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DGPTLRWISFC 16  
 DB 619 NGVTPRRWLSFC 630

## RESULT 13

S00503  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - Pacific electric ray  
 C;Species: Torpedo californica (Pacific electric ray)  
 C;Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
 C;Accession: S00503; S28885; S29880  
 R;Kawakami, K.; Noguchi, S.; Node, M.; Takahashi, H.; Ohta, T.; Kawamura, M.; Nojima, H  
 Nature 316, 733-736, 1995  
 A;Title: Primary structure of the alpha-subunit of Torpedo californica (Na(+)+K(+))ATPa  
 A;Reference number: S00503; PMID:85296307; PMID:2593505

A;Accession: S00503  
 A;Molecule type: mRNA  
 A;Residues: 1-1022 <KAW1>  
 A;Cross-references: UNIPROT:P05025; EMBL:X02810; NID:964399; PIDN:CAA26578.1; PID:96440  
 A;Accession: S28885  
 A;Molecule type: protein  
 A;Residues: 228-240/431-438/535-550/671-690/1011-1022 <KAW2>  
 R;Ohta, T.; Nagano, K.; Yoshida, M.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 2071-2075, 1986

A;Title: The active site structure of Na(+)/K(+)-transporting ATPase: location of the 5  
 A;Reference number: S29880; PMID:86177549; PMID:3008150  
 A;Accession: S29880  
 A;Molecule type: protein  
 A;Residues: 386-402/502-512/671-689/887-906 <OHT>  
 C;Superfamily: Na+/+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C;Keywords: ATP; heterodimer; hydrolyase; ion transport; phosphoprotein; potassium trans  
 F;96-110/Domain: transmembrane #status predicted <TM1>  
 F;130-149/Domain: transmembrane #status predicted <TM2>  
 F;150-290/Domain: intracellular #status predicted <INT2>  
 F;291-313/Domain: transmembrane #status predicted <TM3>  
 F;320-348/Domain: transmembrane #status predicted <TM4>  
 F;349-785/Domain: intracellular #status predicted <INT3>  
 F;586-782/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F;786-809/Domain: transmembrane #status predicted <TM5>  
 F;848-873/Domain: transmembrane #status predicted <TM6>  
 F;874-951/Domain: intracellular #status predicted <INT4>  
 F;952-977/Domain: transmembrane #status predicted <TM7>  
 F;978-1022/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:507/Binding site: ATP (Lys) #status predicted  
 F:716,720,725/Active site: Asp, Asp, Lys #status predicted

Query Match 40.4%; Score 44; DB 1; Length 1022;  
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWISFC 16  
 |||  
 Db 84 PTPPEWIKFC 93

## RESULT 14

A24414  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - human  
 N:Alternate names: sodium pump; sodium/potassium transporting ATPase alpha-A chain  
 C:Species: Homo sapiens (hmn)  
 C>Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 09-Jul-2004  
 C/Accession: A24414; A27795; A39910; I60116; S09171  
 R:Kawakami, K.; Ohta, T.; Nojima, H.; Nagano, K.  
 J. Biochem. 100, 389-397, 1986  
 A>Title: Primary structure of the alpha-subunit of human Na,K-ATPase deduced from cDNA  
 A:Reference number: A24414; MUID:87057096; PMID:2430951  
 A:Accession: A24414  
 A:Molecule type: mRNA  
 A:Residues: 1-1023 <RAW>  
 A/Cross-references: UNIPROT:P05023; EMBL:X04297; NID:g28926; PIDN:CAA27840.1; PID:g28927  
 R:Shull, M.M.; Lingrel, J.B.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
 A>Title: Multiple genes encode the human Na,K-ATPase catalytic subunit.  
 A:Reference number: A94158; MUID:87231946; PMID:3035563  
 A:Accession: A27795  
 A:Molecule type: DNA  
 A:Residues: 168-189,213-214, 'X', 216-244 <SHU>  
 R:Chehab, F.F.; Kan, Y.W.; Law, M.L.; Hartz, J.; Kao, F.T.; Blostein, R.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 7901-7905, 1987  
 A>Title: Human placental Na,K-ATPase alpha subunit: cDNA cloning, tissue expression, D  
 A:Reference number: A39910; MUID:88068506; PMID:2891135  
 A:Accession: A39910  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 199-942 <CHE>  
 A/Cross-references: GB:U03007  
 R:Shull, M.M.; Pugh, D.G.; Lingrel, J.B.  
 Genomics 6, 451-460, 1990  
 A>Title: The human Na,K-ATPase alpha 1 gene: characterization of the 5'-flanking region  
 A:Reference number: I60116; MUID:90228961; PMID:1970326  
 A:Accession: I60116  
 A:Status: translation not shown; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-61 <RES>  
 A/Cross-references: GB:M30310; NID:g179206; PIDN:AAA51801.1; PID:g179208  
 C/Genetics:  
 A:Gene: GDB:ATP1A1  
 A/Cross-references: GDB:119711; OMIM:182310  
 A:Map position: 1p13-1p11  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP, heterodimer; hydrolase; ion transport; osmoregulation; phosphoprotein;  
 F:6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>  
 F:6-95/Domain: intracellular #status predicted <INT1>  
 F:96-120/Domain: transmembrane #status predicted <TM1>  
 F:130-149/Domain: transmembrane #status predicted <TM2>  
 F:150-290/Domain: intracellular #status predicted <INT2>  
 F:291-313/Domain: transmembrane #status predicted <TM4>  
 F:320-348/Domain: transmembrane #status predicted <TM4>  
 F:349-766/Domain: intracellular #status predicted <INT3>  
 F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:787-810/Domain: transmembrane #status predicted <TM5>  
 F:849-874/Domain: transmembrane #status predicted <TM6>  
 F:875-952/Domain: intracellular #status predicted <INT4>  
 F:953-978/Domain: transmembrane #status predicted <TM7>  
 F:979-1023/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:508/Binding site: ATP (Lys) #status predicted  
 F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 40.4%; Score 44; DB 2; Length 1023;  
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWISFC 16  
 |||  
 Db 84 PTPPEWIKFC 93

## RESULT 15

T39685  
 conserved hypothetical protein SPBC1778.03c - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: T39685  
 R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, March 1998  
 A:Reference number: Z21869  
 A:Accession: T39685  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-376 <OLI>  
 A/Cross-references: UNIPROT:Q9Y7J0; EMBL:AL049489; PIDN:CB39798.1; GSPDB:GN00067; SPDB  
 A:Experimental source: strain 972h-; cosmid cl778  
 C/Genetics:  
 A:Gene: SPDB:SPBC1778.03c  
 A:Map position: 2  
 A:introns: 11/2

Query Match 39.9%; Score 43.5; DB 2; Length 376;  
 Best Local Similarity 42.9%; Pred. No. 79;  
 Matches 9; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 1 GGCADGPTLEWIS--FCGG 18  
 |||  
 Db 164 GACAFARSIDWISRYRCPG 184

## RESULT 16

A89813  
 glutamate synthase large subunit [imported] - Staphylococcus aureus (strain N315)  
 C/Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 16-Aug-2004  
 C/Accession: A89813

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu  
 ma, A.; Mizutani-Oli, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.,  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89813

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1499 <RUR>

A/Cross-references: UNIPROT:Q99WD1; GB:BA000018; PID:g13700362; PIDN:BA041660.1; GSPDB

A:Experimental source: strain N315

C/Genetics:

A:Gene: gltB

C:Superfamily: Glutamate synthase, large subunit

Query Match 39.9%; Score 43.5; DB 2; Length 1499;  
 Best Local Similarity 64.3%; Pred. No. 2.8e+02;  
 Matches 9; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 5 DGPTLRWISFCGG 18  
 |||  
 Db 339 DGPTM--ISFCNG 349

```
RESULT 17
D72595
hypothetical protein ABE1229 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: D72595
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A/Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: D72595
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <KAW>
A/Cross-references: UNIPROT:Q9YCM9; DDBJ:AP000061; NID:95104821; PIDN:BAA80218.1; PID:dl
A/Experimental source: strain K1
A/Genetics:
A/Gene: ABE1229

Query Match          39.4%; Score 43; DB 2; Length 113;
Best Local Similarity 61.5%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 GPTLRWISFCG 18
    ||| ||| |||
    21 GEARLRCWPSRCRG 33

RESULT 18
T15386
hypothetical protein C03B1.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T15386
R/Martin, J.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid C03B1.
A/Reference number: Z18340
A/Accession: T15386
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-115 <MAR>
A/Cross-references: UNIPROT:Q11110; EMBL:U04952; NID:91072237; PID:91072244; PIDN:AAA817
C/Genetics:
A/Gene: CESP:C03B1.3
A/Introns: 80/1

Query Match          39.4%; Score 43; DB 2; Length 115;
Best Local Similarity 46.7%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 3 CADGPTLRWISFCG 17
    ||| ||| |||
    68 CASGEVHYHWACFCG 82

RESULT 19
I48685
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I48685; S43172
R/Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A>Title: Genes for mast-cell serine protease and their molecular evolution.
A/Reference number: I48684; MUID:95048582; PMID:7959952
A/Accession: I48685
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-230 <RES>
A/Cross-references: UNIPROT:P21845; EMBL:X78542; NID:9468809; PIDN:CAA55288.1; PID:94688
C/Superfamily: trypsin; trypsin homology
```

```
C/Keywords: hydrolase; serine proteinase
F:32-230/Domain: trypsin homology #status atypical <TRY>

Query Match          39.4%; Score 43; DB 2; Length 230;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 9 LBEWISFCG 18
    ||| ||| |||
    53 LNWYHIFCG 62

RESULT 20
AB2768
lipase biosynthesis protein B [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AB2768
R/Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.E.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AB2768
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-233 <KUR>
A/Cross-references: UNIPROT:Q8UF44; GB:AE008688; PIDN:AAL42560.1; PID:917739983; GSPDB:1
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: lipB
A/Map position: circular chromosome
C/Superfamily: Escherichia coli lipase-protein ligase lipB

Query Match          39.4%; Score 43; DB 2; Length 233;
Best Local Similarity 43.5%; Pred. No. 60;
Matches 10; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

OY 1 GGCAD-----GPTLRWISFCG 17
    ||| : ||| : ||| : |||
    148 GGAEDKIALALGIRLKNVSYFHG 170

RESULT 21
T19988
hypothetical protein C47B2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19988
R/Kershaw, J.
submitted to the EMBL Data Library, October 1997
A/Reference number: Z19208
A/Accession: T19988
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-246 <WIL>
A/Cross-references: UNIPROT:Q62106; EMBL:Z99709; PIDN:CAB16860.1; GSPDB:GN00019; CESP:C
C/Genetics:
A/Gene: CESP:C47B2.5
A/Map position: 1
A/Introns: 91/3; 127/3
C/Superfamily: conserved hypothetical protein YP016c

Query Match          39.4%; Score 43; DB 2; Length 246;
Best Local Similarity 41.7%; Pred. No. 63;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 6 GPTLRWISFCG 17
    ||| : ||| : |||
```





A:Experimental source: strain AB972  
 C:Genetic8  
 A:Gene: SGD:RSN1; MIPS:YMR266w  
 A:Cross-references: SGD:S0004879  
 A:Map position: 13R  
 C:Superfamily: yeast probable membrane protein Y0I084w  
 C:Keywords: transmembrane protein  
 F:12-48/Domain: transmembrane #status predicted <TM1>  
 F:106-122/Domain: transmembrane #status predicted <TM2>  
 F:152-168/Domain: transmembrane #status predicted <TM3>  
 F:195-411/Domain: transmembrane #status predicted <TM4>  
 F:335-451/Domain: transmembrane #status predicted <TM5>  
 F:545-561/Domain: transmembrane #status predicted <TM6>  
 F:599-615/Domain: transmembrane #status predicted <TM7>  
 F:646-662/Domain: transmembrane #status predicted <TM8>  
 F:668-684/Domain: transmembrane #status predicted <TM9>

Query Match 39.4%; Score 43; DB 2; Length 953;  
 Best Local Similarity 41.2%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 GCGADGPTLRWISFC 17  
 DB 558 GAFIDGTVRKRMKRFCS 574

RESULT 26  
 B37227  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 09-Jul-2004  
 C:Accession: B37227; I50395  
 R:Takeyasu, K.; Lemas, V.; Fambrough, D.M.  
 Am. J. Physiol. 259, C619-C630, 1990  
 A:Title: Stability of Na(+)-K(+) ATPase alpha-subunit isoforms in evolution.  
 A:Reference number: A37227; PMID:91023019; PMID:2171348  
 A:Accession: B37227  
 A:Molecule type: mRNA  
 A:Residues: 1-1010 <TA2>  
 A:Cross-references: UNIPROT:P24799; GB:M59960; NID:G212407; PID:AAA48982.1; PID:G212408  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; sodium tr  
 F:574-770/Domain: ATPase nucleotide-binding domain homology <ATP>  
 F:202-470/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:363/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:495/Binding site: ATP (Lys) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 1010;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLRWISFC 16  
 DB 71 PTPPEWVKFC 80

RESULT 27  
 S00801  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - human  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
 C:Accession: S00801; S04019; A27397; S02275  
 R:Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Broude, N.E.; Ushkarev, Y.A.; Melkov, A.M.; S  
 dyanov, N.N.; Sverdlov, E.D.  
 FEBS Lett. 233, 87-94, 1988  
 A:Title: Family of human Na,K-ATPase genes. Structure of the gene for the catalytic subu  
 A:Reference number: S00801; MUID:88255304; PMID:2838329  
 A:Accession: S00801  
 A:Molecule type: DNA  
 A:Residues: 1-1013 <OV>  
 A:Cross-references: UNIPROT:P13637; EMBL:M37456  
 R:Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkarev, Y.A.; Melkov, A.M.; Smir  
 ov, N.N.; Ovchinnikov, Y.A.

Dokl. Biochem. 297, 426-431, 1987  
 A:Title: Family of human Na(+),K(+)-ATPase genes. Structure of the gene of isoform alph  
 A:Reference number: S04019  
 A:Accession: S04019  
 A:Molecule type: DNA  
 A:Residues: 1, 'RHH', 3-1013 <SVE1>  
 A:Cross-references: EMBL:X12910; NID:G28963  
 A>Note: the authors translated the codon TTC for residue 283 as Ser and TCT for residue  
 A>Note: this paper is a translation of the Russian paper published in Dokl. Akad. Nauk;  
 R:Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkarev, Y.A.; Alimkete, R.L.; I  
 tina, M.B.; Sverdlov, V.E.; Modyanov, N.N.; Ovchinnikov, Y.A.  
 FEBS Lett. 217, 275-278, 1987  
 A:Title: The family of human Na+, K+-ATPase genes. No less than five genes and/or pseudo  
 A:Reference number: A27397; MUID:87247232; PMID:3036582  
 A:Accession: A27397  
 A:Molecule type: mRNA  
 A:Residues: 243-434 <SVE2>  
 A:Cross-references: GB:M27570  
 A:Gene: GDB:ATP1A3  
 A:Cross-references: GDB:119713; OMIM:182350  
 A:Map position: 19q13.2-19q13.2  
 A:Introns: 2/3; 31/3; 51/3; 119/3; 157/3; 202/3; 242/1; 331/3; 398/1; 435/2; 479/3; 544/  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
 F:86-110/Domain: transmembrane #status predicted <TM1>  
 F:120-139/Domain: transmembrane #status predicted <TM2>  
 F:140-280/Domain: intracellular #status predicted <TM3>  
 F:281-303/Domain: transmembrane #status predicted <TM4>  
 F:310-338/Domain: transmembrane #status predicted <TM5>  
 F:339-776/Domain: intracellular #status predicted <TM3>  
 F:577-773/Domain: ATPase nucleotide-binding domain homology <ATP>  
 F:777-800/Domain: transmembrane #status predicted <TM5>  
 F:839-864/Domain: transmembrane #status predicted <TM6>  
 F:865-942/Domain: intracellular #status predicted <TM7>  
 F:943-968/Domain: transmembrane #status predicted <TM7>  
 F:969-1013/Domain: extracellular #status predicted <EXT>  
 F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:498/Binding site: ATP (Lys) #status predicted  
 F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match 39.4%; Score 43; DB 1; Length 1013;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLRWISFC 16  
 DB 74 PTPPEWVKFC 83

RESULT 28  
 C24639  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - rat  
 N:Alternate names: Na+/K+-transporting ATPase alpha (III) chain  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 30-Jun-1988 #sequence\_revision 23-Apr-1993 #text\_change 09-Jul-2004  
 C:Accession: C24639; S00514; B27180; A60470  
 R:Shull, G.E.; Greeb, J.; Lingrel, J.B.  
 Biochemistry 25, 8125-8132, 1986  
 A:Title: Molecular cloning of three distinct forms of the Na+,K+-ATPase alpha-subunit f  
 A:Reference number: A90512; MUID:87128908; PMID:3028470  
 A:Accession: C24639  
 A:Molecule type: mRNA  
 A:Residues: 1-1013 <SHU>  
 A:Cross-references: UNIPROT:P06687; EMBL:M14513; NID:G203030; PID:AAA40777.1; PID:G203  
 A>Note: in the authors' translation 405-Ser is shown after residue 409 and, consequentl  
 R:Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamine, H.; Ohta, T.;  
 U. Biochem. 102, 43-58, 1987  
 A:Title: Primary structures of two types of alpha-subunit of rat brain Na(+),K(+)-ATP  
 A:Reference number: S00460; MUID:88032933; PMID:2822682  
 A:Accession: S00514  
 A:Molecule type: mRNA  
 A:Residues: 1-907, 'C', 909-1013 <HAR>

A;Cross-references: EMBL:X05883; NID:955769; PDB:CA29307.1; PID:955770  
 R;Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.  
 J. Cell Biol. 105, 1855-1865, 1987  
 A;Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural  
 A;Reference number: A92749; MUID:8803255; PMID:2822726  
 A;Accession: B27180  
 A;Molecule type: mRNA  
 A;Residues: 1,'NL','4'-103,'R','105-113,'E','115-127,'G','129-148,'Q','150-151,'T','153-165,'D'  
 A;Cross-references: EMBL:M8648; NID:9205633; PDB:AAA1672.1; PID:9205634  
 A;Note: the authors translated the codon CAG for residue 149 as Glu, GGC for residue 194  
 R;Hsu, Y.M.; Guidotti, G.  
 Biochemistry 28, 569-573, 1989  
 A;Title: Rat brain has the alpha3 form of the (Na,K)-ATPase.  
 A;Reference number: A60470; MUID:89229049; PMID:2540801  
 A;Accession: A60470  
 A;Molecule type: protein  
 A;Residues: 117-132;586-595,'X','597-601 <HSU>  
 A;Comment: The alpha-3 form appears to be highly ouabain-inhibitable, as is alpha-2 but  
 C;Genetics:  
 A;Gene: NKAA3  
 C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F;86-110/Domain: transmembrane #status predicted <TM1>  
 F;120-139/Domain: transmembrane #status predicted <TM2>  
 F;140-280/Domain: intracellular #status predicted <INT7>  
 F;281-303/Domain: transmembrane #status predicted <TM3>  
 F;310-338/Domain: transmembrane #status predicted <TM4>  
 F;339-776/Domain: intracellular #status predicted <INT3>  
 F;577-773/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F;777-800/Domain: transmembrane #status predicted <TM5>  
 F;839-864/Domain: transmembrane #status predicted <TM6>  
 F;865-942/Domain: intracellular #status predicted <INT4>  
 F;943-968/Domain: transmembrane #status predicted <TM7>  
 F;969-1013/Domain: extracellular #status predicted <EXT>  
 F;366/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F;448/Binding site: ATP (Lys) #status predicted  
 F;707,'711,'716/Active site: Asp, Asp, Lys #status predicted  
 Query Match 39.4%; Score 43; DB 2; Length 1013;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 7 PTLREWISFC 16  
 DB 74 PTLPEWVKFC 83  
 RESULT 29  
 A37227  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - chicken  
 C;Species: Gallus gallus (chicken)  
 C;Date: 16-Sep-1992 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C;Accession: I50394; A37227  
 R;Takeyasu, K.; Lemas, M.; Fambrough, D.M.  
 Am. J. Physiol. 259, 619-630, 1991  
 A;Title: Stability of the Na+,K+-ATPase alpha-subunit isoforms in evolution.  
 A;Reference number: I50394  
 A;Accession: I50394  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1017 <TMX>  
 A;Cross-references: UNIPROT:P24797; GB:M59595; NID:9212405; PDB:AAA48981.1; PID:9212406  
 R;Takeyasu, K.; Lemas, V.; Fambrough, D.M.  
 Am. J. Physiol. 259, C619-C630, 1990  
 A;Title: Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.  
 A;Reference number: A37227; MUID:91023019; PMID:2171348  
 A;Accession: A37227  
 A;Molecule type: mRNA  
 A;Residues: 3-1017 <TA2>  
 C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C;Keywords: ATP; glycoprotein; hydrolase; phosphoprotein  
 F;581-777/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F;210,478/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;371/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 Query Match 39.4%; Score 43; DB 2; Length 1017;  
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 7 PTLREWISFC 16  
 DB 79 PTLPEWVKFC 88  
 RESULT 30  
 A34474  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - human  
 A;Alternate names: Na+/K+-exchanging ATPase alpha chain-4; sodium/potassium transportin  
 C;Species: Homo sapiens (man)  
 C;Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 09-Jul-2004  
 C;Accession: A34474; B27795; D27397  
 R;Shull, M.M.; Pugh, D.G.; Lingrel, J.B.  
 J. Biol. Chem. 264, 17532-17543, 1989  
 A;Title: Characterization of the human Na,K-ATPase alpha2 gene and identification of in  
 A;Reference number: A34474; MUID:9008924; PMID:2477373  
 A;Accession: A34474  
 A;Molecule type: DNA  
 A;Residues: 1-1020 <SHU>  
 A;Cross-references: UNIPROT:P50993; GB:J05096; NID:9179164; PDB:AAA51797.1; PID:917916  
 R;Shull, M.M.; Lingrel, J.B.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
 A;Title: Multiple genes encode the human Na+,K+-ATPase catalytic subunit.  
 A;Reference number: A94158; MUID:87231946; PMID:3035563  
 A;Accession: B27795  
 A;Molecule type: DNA  
 A;Residues: 211-249 <SH2>  
 A;Cross-references: GB:M16795; NID:9179196; PDB:AAA51799.1; PID:9553194  
 R;Sverdlov, E.D.; Monastyrskaya, G.S.; Brode, N.E.; Ushkaryov, Y.A.; Allikmeets, R.L.;  
 Lina, M.B.; Sverdlov, V.E.; Modyanov, N.N.; Ovchinnikov, Y.A.  
 FEBS Lett. 217, 275-278, 1987  
 A;Title: The family of human Na+,K+-ATPase genes. No less than five genes and/or pseudo  
 A;Reference number: A27397; MUID:87247232; PMID:3036582  
 A;Accession: D27397  
 A;Molecule type: DNA  
 A;Residues: 251-442 <SVB>  
 A;Cross-references: GB:M27571  
 C;Genetics:  
 A;Gene: GDB:ATP1A2  
 A;Cross-references: GDB:119712; OMIM:182340  
 A;Map position: 1q21-q23  
 C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
 F;6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MAT>  
 F;6-93/Domain: intracellular #status predicted <INT1>  
 F;94-118/Domain: transmembrane #status predicted <TM1>  
 F;112-147/Domain: transmembrane #status predicted <TM2>  
 F;148-288/Domain: intracellular #status predicted <INT3>  
 F;289-311/Domain: transmembrane #status predicted <TM3>  
 F;318-346/Domain: transmembrane #status predicted <TM4>  
 F;347-783/Domain: intracellular #status predicted <INT4>  
 F;584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F;784-807/Domain: transmembrane #status predicted <TM5>  
 F;846-871/Domain: transmembrane #status predicted <TM6>  
 F;872-949/Domain: intracellular #status predicted <INT5>  
 F;950-975/Domain: transmembrane #status predicted <TM7>  
 F;976-1020/Domain: extracellular #status predicted <EXT>  
 F;376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F;505/Binding site: ATP (Lys) #status predicted  
 F;714,718,723/Active site: Asp, Asp, Lys #status predicted  
 Query Match 39.4%; Score 43; DB 2; Length 1020;  
 Best Local Similarity 60.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 7 PTLREWISFC 16

Db 82 PTPPEWVKFC 91

## RESULT 31

B24639

Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha-2 chain - rat  
N/Alternate names: Na+/K+-transporting ATPase alpha-plus chain

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004

C/Accession: B24639

R/Shull, G.E.; Greep, J.; Lingrel, J.B.

Biochemistry 25, 8125-8132, 1986

A/Title: Molecular cloning of three distinct forms of the Na<sup>+</sup>, K<sup>+</sup>-ATPase alpha-subunit from

A/Reference number: A90512; MUID:87128908; PMID:3028470

A/Accession: B24639

A/Molecule type: mRNA

A/Residues: 1-1020 <SHU>

A/Cross-references: UNIPROT:P06686; EMBL:M14512; NID:G203028; PIDN:AAA40776.1; PID:G2030

C/Genetics:

A/Gene: NKA2

C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

F/6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MAT>

F/6-93/Domain: intracellular #status predicted <INT1>

F/94-119/Domain: transmembrane #status predicted <TM1>

F/128-147/Domain: transmembrane #status predicted <TM2>

F/148-288/Domain: intracellular #status predicted <INT2>

F/289-311/Domain: transmembrane #status predicted <TM3>

F/318-346/Domain: intracellular #status predicted <INT3>

F/347-783/Domain: intracellular #status predicted <INT3>

F/584-780/Domain: ATPase nucleotide-binding domain homology <ATN>

F/784-807/Domain: transmembrane #status predicted <TM5>

F/846-871/Domain: transmembrane #status predicted <TM6>

F/872-949/Domain: intracellular #status predicted <INT4>

F/950-975/Domain: transmembrane #status predicted <TM7>

F/976-1020/Domain: extracellular #status predicted <EXT>

F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted

F/505/Binding site: ATP (Lys) #status predicted

F/714,718,723/Active site: Asp, Asp, Lys #status predicted

Query Match 39.4%; Score 43; DB 2; Length 1020;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTPPEWVKFC 16

Db 82 PTPPEWVKFC 91

## RESULT 32

PWSHNA

Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha chain precursor - sheep

N/Alternate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha chain

C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004

C/Accession: A01074; A35426

R/Shull, G.E.; Schwartz, A.; Lingrel, J.B.

Nature 316, 691-695, 1985

A/Title: Amino-acid sequence of the catalytic subunit of the (Na<sup>+</sup>)+K<sup>+</sup>(+) ATPase deduced

A/Reference number: A01074; MUID:85296299; PMID:2993903

A/Accession: A01074

A/Molecule type: mRNA

A/Residues: 1-1021 <SHU>

A/Cross-references: UNIPROT:P04074; GB:X02813; NID:G1205; PIDN:CAA36581.1; PID:G1206

R/Hinz, H.R.; Kirsley, T.L.

J. Biol. Chem. 265, 10260-10265, 1990

A/Title: Lysine 480 is an essential residue in the putative ATP site of lamb kidney (Na<sup>+</sup>,

A/Reference number: A35426; MUID:90285144; PMID:2162343

A/Accession: A35426

A/Status: preliminary

A/Molecule type: protein

A/Residues: 475-492 <HN>

C/Comment: This is the catalytic component of the active enzyme, which catalyzes the hyd

reates the electrochemical gradient of sodium and potassium, providing the energy for a

n function.

C/Comment: This enzyme is specifically inhibited by cardiac glycosides such as digoxin.

C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C/Keywords: ATP; hydrolase; phosphoprotein; potassium transport; sodium transport; tran

F/6-1021/Product: Na+/K+-transporting ATPase alpha chain #status predicted <MAT>

F/94-115/Domain: transmembrane #status predicted <TM1>

F/128-144/Domain: transmembrane #status predicted <TM2>

F/289-311/Domain: transmembrane #status predicted <TM3>

F/318-346/Domain: transmembrane #status predicted <TM4>

F/585-781/Domain: ATPase nucleotide-binding domain homology <ATN>

F/785-808/Domain: transmembrane #status predicted <TM5>

F/947-872/Domain: transmembrane #status predicted <TM6>

F/951-976/Domain: transmembrane #status predicted <TM7>

F/315/Binding site: cardiac glycoside (TTP) #status predicted

F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted

F/506/Binding site: ATP (Lys) #status predicted

Query Match 39.4%; Score 43; DB 1; Length 1021;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTPPEWVKFC 16

Db 82 PTPPEWVKFC 91

## RESULT 33

S04630

Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha-1 chain - horse

C/Species: Equus caballus (domestic horse)

C/Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004

C/Accession: S04630

R/Kano, I.; Nagai, F.; Satoh, K.; Uehiyama, K.; Nakao, T.; Kano, K.

FEBS Lett. 250, 91-98, 1989

A/Title: Structure of the alpha(1) subunit of horse Na,K-ATPase gene.

A/Reference number: S04630; MUID:89290042; PMID:2544461

A/Accession: S04630

A/Molecule type: DNA

A/Residues: 1-1021 <KAN>

A/Cross-references: UNIPROT:P18907; EMBL:X16773; NID:G1010; PIDN:CAA34716.1; PID:G87102

C/Genetics:

A/Intons: 4/3; 39/3; 59/3; 127/3; 165/3; 210/3; 250/1; 339/3; 406/1; 442/3; 487/3; 552;

C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium tran

F/6-1021/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>

F/6-93/Domain: intracellular #status predicted <INT1>

F/94-118/Domain: transmembrane #status predicted <TM1>

F/128-147/Domain: transmembrane #status predicted <TM2>

F/148-288/Domain: intracellular #status predicted <INT2>

F/289-311/Domain: transmembrane #status predicted <TM3>

F/318-346/Domain: transmembrane #status predicted <TM4>

F/347-784/Domain: intracellular #status predicted <INT3>

F/585-781/Domain: ATPase nucleotide-binding domain homology <ATN>

F/785-808/Domain: transmembrane #status predicted <TM5>

F/847-872/Domain: transmembrane #status predicted <TM6>

F/873-950/Domain: intracellular #status predicted <INT4>

F/971-976/Domain: transmembrane #status predicted <TM7>

F/977-1021/Domain: extracellular #status predicted <EXT>

F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted

F/506/Binding site: ATP (Lys) #status predicted

F/715,719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 39.4%; Score 43; DB 1; Length 1021;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTPPEWVKFC 16

Db 82 PTPPEWVKFC 91

## RESULT 34

A28199  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 21-Sep-1988 #sequence\_revision 21-Sep-1988 #text\_change 09-Jul-2004  
C:Accession: A28199  
R:Takeyasu, K.; Tamkun, M.M.; Renaud, K.J.; Fambrough, D.M.  
J. Biol. Chem. 263, 4347-4354, 1988  
A:Title: Ouabain-sensitive (Na(+)+K(+))-ATPase activity expressed in mouse L cells by  
A:Reference number: A28199; MUID:88153759; PMID:2831227  
A:Accession: A28199  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1021 <TAK>  
A:Cross-references: UNIPROT:P09572; GB:J03230; NID:9211219; PIDN:AAA48607.1; PID:9211220240  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein  
E:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:213,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:506/Binding site: ATP (lys) #status predicted

Query Match	39.4%	Score 43	DB 2	Length 1021
Best Local Similarity	60.0%	Pred. No. 2.4e+02		
Matches	6	Conservative	1	Mismatches 3
				Indels 0
				Gaps 0
Qy	7	PTLRWISFC	16	
Db	82	PTTPWVKFC	91	

RESULT 35  
 B24862  
 Na+/K+-exchanging ATPase (EC 3. 6. 3. 9) alpha chain - pig  
 N:Alternate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha chain  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 30-Jun-1988 #sequence: revision 30-Jun-1988 #text\_change 09-Jul-2004  
 C:Accession: B24862; 146572; A35504; 500011; 500502; S02569; S29762  
 R:Ovchinnikov, Y.A.; Modyanov, N.N.; Brode, N.E.; Petrkhin, K.E.; Grishin, A.V.; Arzam  
 FEBS Lett. 201, 237-245, 1986  
 A:Title: Pig kidney Na+, K+-ATPase. Primary structure and spatial organization.  
 A:Reference number: A91361; MUID:86220813; PMID:2423371  
 A:Accession: B24862  
 A:Molecule type: mRNA  
 A:Residues: 1-1021 <OVCA>  
 A:Cross-references: UNIPROT:P05024; EMBL:X03938; NID:g16897; PIDN:CAA27576.1; PID:g16898  
 A:Note: the authors translated the codon TCC for residue 391 as Phe, TCG for residue 722  
 A:Note: part of this sequence, including the amino and carboxyl end of the mature protei  
 R:Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Arsenyan, S.G.; Brode, N.E.; Petrkhin, K.E.;  
 Dokl. Biochem. 283, 270-272, 1985  
 A:Title: Amino acid sequence of the 17-kilodalton fragment of the cytoplasmic region of  
 A:Reference number: I46572  
 A:Accession: I46572  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 469-617 <OVCA>  
 A:Cross-references: GB:X32512; NID:g164385; PIDN:AAA1004.1; PID:g164386  
 R:Karlish, S.J.D.; Goldshleger, R.; Stein, W.D.  
 P:Karlsh. Acad. Sci. U.S.A. 87, 4566-4570, 1990  
 A:Title: A 19-kDa C-terminal tryptic fragment of the alpha chain of Na/K-ATPase is essent  
 A:Reference number: A35504; MUID:90280416; PMID:2162048  
 A:Accession: A35504  
 A:Molecule type: Protein  
 A:Residues: 836-845 'R', 847-851 <KAR>  
 R:Ovchinnikov, Y.A.; Arzamazova, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Aldanova, N  
 FEBS Lett. 217, 269-274, 1987  
 A:Title: Detailed structural analysis of exposed domains of membrane-bound Na+, K+-ATPase  
 A:Accession number: S00011; MUID:87247231; PMID:3036581  
 A:Contents: annotation; membrane topology  
 R:Ovchinnikov, Y.A.; Luneva, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Arzamazova, N.M.  
 FEBS Lett. 227, 230-234, 1988  
 A:Title: Topology of Na+, K-ATPase: identification of the extra- and intracellular hydroph  
 A:Reference number: S02569; MUID:86112852; PMID:2448169  
 A:Contents: annotation; membrane topology

C:Superfamily:aa/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP, heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
 F:6-1021/Product: Na+/K+-transporting ATPase alpha chain #status experimental <Mat>  
 F:6-93/Domain: intracellular #status predicted <INT1>  
 F:94-118/Domain: transmembrane #status predicted <TM1>  
 F:128-147/Domain: transmembrane #status predicted <TM2>  
 F:148-288/Domain: intracellular #status predicted <INT2>  
 F:289-311/Domain: transmembrane #status predicted <TM3>  
 F:318-346/Domain: transmembrane #status predicted <TM4>  
 F:347-784/Domain: intracellular #status predicted <INT3>  
 F:568-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:785-808/Domain: transmembrane #status predicted <TM5>  
 F:847-872/Domain: transmembrane #status predicted <TM6>  
 F:873-950/Domain: intracellular #status predicted <INT4>  
 F:951-976/Domain: transmembrane #status predicted <TM7>  
 F:977-1021/Domain: extracellular #status predicted <EXT>  
 F:137/Active site: ATP (aspartylphosphate intermediate) #status predicted  
 F:1506/Binding site: Asp (Lys) #status predicted  
 F:715, 719, 724/Active site: Asp, Asp, Lys #status predicted

Query Match	39.4%	Score 43;	DB 2;	Length 1021;
Best Local Similarity	60.0%	Pred. NO. 2.4e+02;		
Matches	6;	Conservative	1;	Mismatches 5;
				Indels 0;
				Gaps 0;
Qy	7	PTLRWISFC	16	
Db	82	PTLRWISFC	91	

RESULT 36  
S49127  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - European eel  
C|Species: *Anguilla anguilla* (European eel)  
C|Date: 01-Feb-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004  
C|Accession: S49127  
R|Cutter: C.; Sanders, I.L.; Cramb, G.  
submitted to the EMBL Data Library, November 1993  
A|Reference number: S45093  
A|Accession: S49127  
A|Status: preliminary  
A|Molecule type: mRNA  
A|Residues: 1-1022 <CUT>  
A|Cross-references: UNIPROT:Q92030; EMBL:X76108; NID:G509405; PIRN:CAA53714.1; PIR:G509405  
C|Superfamily: Na+/K+-transporting ATPase alpha chain, ATPase nucleotide-binding domain  
C|Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; transmembrane  
F|568-782/Domain: ATPase nucleotide-binding domain homology <ATN>  
F|214,483/Binding site: carboxylate (Asn) (covalent) #status predicted  
F|375/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F|507/Binding site: ATP (lyse) #status predicted

Query Match	39.4%	Score 43;	DB 2;	Length 1022;
Best Local Similarity	60.0%	Pred. No. 2.4e+02;		
Matches	6;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0

RESULT 37  
 A24639  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain [validated] - rat  
 N:Alternate names: Na+/K+-transporting ATPase alpha chain, kidney-type  
 N:Contents: Na+/K+-transporting ATPase alpha-S chain  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 18-Aug-2000 #sequence revision 18-Aug-2000 #text-change 09-Jul-2004  
 C:Accession: A24639; S00460; A27180; S11020; A25171; S29877; S10758  
 R:Smith, G.E.; Greeb, J.; Lingrel, J.B.  
 Biochemistry 25, 8125-8132, 1986  
 A:Title: Molecular cloning of three distinct forms of the Na+,K+-ATPase alpha-subunit f  
 A:Reference number: A90512; MUID:87128908; PMID:3028470  
 A:Accession: A24639  
 A:Molecule type: mRNA

A;Residues: 1-1023 <SHU>  
A;Cross-references: UNIPROT:P06685; EMBL:M4511; NID:G203026; PIDN:AAA40775.1; PID:G203026  
R;Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohra, T.; N  
J. Biochem. 102, 43-58, 1987  
A;Title: Primary structures of two types of alpha-subunit of rat brain Na(+)/K(+) -ATPase  
A;Reference number: S00460; MUID:88032933; PMID:2822682  
A;Accession: S00460  
A;Molecule type: mRNA  
A;Residues: 1-1023 <HAR>  
A;Cross-references: EMBL:X05882; NID:G55771; PIDN:CAA29306.1; PID:G55772  
R;Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.  
J. Cell Biol. 105, 1855-1865, 1987  
A;Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural a  
A;Reference number: A92749; MUID:88033255; PMID:2822726  
A;Accession: A27180  
A;Molecule type: mRNA  
A;Residues: 1-67, 'PV', 70-174, 'E', 176-187, 'V', 189-334, 'V', 336-1023 <HER>  
A;Cross-references: EMBL:M08647; NID:G205631; PIDN:AAA41671.1; PID:G205632  
R;Nagawa, Y.; Kawakami, K.; Nagano, K.  
Biochim. Biophys. Acta 1049, 286-292, 1990  
A;Title: Cloning and analysis of the 5'-flanking region of rat Na(+)/K(+) -ATPase alpha-1  
A;Reference number: S11020; MUID:90344872; PMID:2166579  
A;Accession: S11020  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-41 <YAG>  
A;Cross-references: EMBL:X53233  
R;Schneider, J.W.; Mercer, R.W.; Caplan, M.; Emanuel, J.R.; Sweadner, K.J.; Benz Jr., E.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6357-6361, 1985  
A;Title: Molecular cloning of rat brain Na,K-ATPase alpha-subunit cDNA.  
A;Reference number: A25171; MUID:85298352; PMID:2994074  
A;Accession: A25171  
A;Molecule type: mRNA  
A;Residues: 489-533 <SCH>  
R;Lytton, J.  
Biochem. Biophys. Res. Commun. 132, 764-769, 1985  
A;Title: The catalytic subunits of the (Na(+), K(+)) -ATPase alpha and alpha(+) isozymes  
A;Reference number: S29877; MUID:86050667; PMID:2998384  
A;Accession: S29877  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 6-19 <LTV>  
R;Kurihara, K.; Hosoi, K.; Kodama, A.; Ueha, T.  
Biochim. Biophys. Acta 1039, 234-240, 1990  
A;Title: A new electrophoretic variant of alpha subunit of Na(+)/K(+) -ATPase from the su  
A;Reference number: S10758; MUID:90304196; PMID:2163680  
A;Accession: S10758  
A;Molecule type: protein  
A;Residues: 6, 'X', 8-10, 'X', 12-16 <KUR>  
A;Experimental source: submandibular gland  
A;Note: designated alpha-S form; thought to arise from alpha-1 chain by post-translation  
C;Genetics:  
A;Gene: NKXAL  
A;Intons: 4/3  
A;Note: the list of introns may be incomplete  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
F;6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status experimental <MAT>  
F;6-95/Domain: intracellular #status predicted <INT1>  
F;96-10/Domain: transmembrane #status predicted <TM1>  
F;110-149/Domain: transmembrane #status predicted <TM2>  
F;150-290/Domain: intracellular #status predicted <INT2>  
F;291-313/Domain: transmembrane #status predicted <TM3>  
F;320-348/Domain: transmembrane #status predicted <TM4>  
F;349-786/Domain: intracellular #status predicted <INT3>  
F;587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;787-810/Domain: transmembrane #status predicted <TM5>  
F;849-874/Domain: transmembrane #status predicted <TM6>  
F;875-952/Domain: intracellular #status predicted <INT4>  
F;953-978/Domain: transmembrane #status predicted <TM7>  
F;979-1023/Domain: extracellular #status predicted <EXT>  
F;376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;508/Binding site: ATP (Lys) #status predicted

F;717,721,726/Active site: Asp, Asp, Lys #status predicted  
Query Match 39.4%; Score 43; DB 1; Length 1023;  
Best Local Similarity 60.0%; Pred. No. 2,4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 7 PTLREWISFC 16  
DB 84 PTLPEWVKFC 93  
RESULT 38  
S24650  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - giant toad  
C;Species: Bufo marinus (giant toad)  
C;Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #ext\_change 09-Jul-2004  
C;Accession: A43451; S24650  
R;Baisner, F.; Canessa, C.M.; Horisberger, J.D.; Rossier, B.C.  
J. Biol. Chem. 267, 16895-16903, 1992  
A;Title: Primary sequence and functional expression of a novel ouabain-resistant Na,K-A  
A;Reference number: A43451; MUID:92380991; PMID:1380956  
A;Accession: A43451  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1023 <JAI>  
A;Cross-references: UNIPROT:P30714; EMBL:Z11798; NID:G62491; PIDN:CAA77842.1; PID:G6249  
A;Experimental source: urinary bladder cell line TBM 18-23  
A;Note: submitted to the EMBL Data Library, March 1992  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
F;6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>  
F;6-95/Domain: intracellular #status predicted <INT1>  
F;96-120/Domain: transmembrane #status predicted <TM1>  
F;130-149/Domain: transmembrane #status predicted <TM2>  
F;150-290/Domain: intracellular #status predicted <INT2>  
F;291-313/Domain: transmembrane #status predicted <TM3>  
F;320-348/Domain: transmembrane #status predicted <TM4>  
F;349-786/Domain: intracellular #status predicted <INT3>  
F;587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;787-810/Domain: transmembrane #status predicted <TM5>  
F;849-874/Domain: transmembrane #status predicted <TM6>  
F;875-952/Domain: intracellular #status predicted <INT4>  
F;953-978/Domain: transmembrane #status predicted <TM7>  
F;979-1023/Domain: extracellular #status predicted <EXT>  
F;376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;508/Binding site: ATP (Lys) #status predicted  
F;717,721,726/Active site: Asp, Asp, Lys #status predicted  
Query Match 39.4%; Score 43; DB 1; Length 1023;  
Best Local Similarity 60.0%; Pred. No. 2,4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 7 PTLREWISFC 16  
DB 84 PTLPEWVKFC 93  
RESULT 39  
A60444  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain precursor - African clawed frog  
N;Alternate names: sodium pump alpha chain  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #ext\_change 09-Jul-2004  
C;Accession: A60444  
R;Verrey, F.; Kairouz, P.; Schaefer, E.; Fuentes, P.; Geering, P.; Rossier, B.C.; Kraeh  
Am. J. Physiol. 256, F1034-F1043, 1989  
A;Title: Primary sequence of Xenopus laevis Na(+)-K(+) -ATPase and its localization in A  
A;Reference number: A60444; MUID:89285429; PMID:2544104  
A;Accession: A60444  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-1025 <VER>



A:Cross-references: UNIPROT:P25489; EMBL:X58629; NID:962641; PIDN:CAA41483.1; PID:962642  
 C:Comment: The alpha chain is the catalytic chain.  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; sodium tr  
 F:589-785/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:217-485/Binding site: carboxylate-binding domain homology <ATN>  
 F:378/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:510/Binding site: Asp (Lys) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 1025;  
 Best Local Similarity 60.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLRWISFC 16  
 DB 86 PTLRWKFC 95

## RESULT 40

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - white sucker  
 C:Species: Catostomus commersoni (white sucker)  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 C:Accession: S14740  
 R:Schneerock, C.; Morley, S.D.; Okawara, Y.; Lederis, K.; Richter, D.  
 Biol. Chem. Hoppe-Seyler 372, 279-286, 1991  
 A:Title: Sodium and potassium ATPase of the teleost fish *Catostomus commersoni*. Sequence  
 A:Reference number: S14740; MUID:91282983; PMID:1711856  
 A:Accession: S14740  
 A:Molecule type: mRNA  
 A:Residues: 1-1027 <SCH>

A:Cross-references: UNIPROT:P25489; EMBL:X58629; NID:962641; PIDN:CAA41483.1; PID:962642  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; hydrolase; ion transport; phosphoprotein; potassium transport; sodium t  
 F:99-124/Domain: transmembrane #status predicted <TM1>  
 F:113-152/Domain: transmembrane #status predicted <INT2>  
 F:153-293/Domain: intracellular #status predicted <INT3>  
 F:294-316/Domain: transmembrane #status predicted <TM3>  
 F:323-351/Domain: transmembrane #status predicted <TM4>  
 F:352-790/Domain: intracellular #status predicted <INT3>  
 F:591-787/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:791-814/Domain: transmembrane #status predicted <TM5>  
 F:853-878/Domain: transmembrane #status predicted <TM6>  
 F:879-956/Domain: intracellular #status predicted <INT4>  
 F:957-982/Domain: transmembrane #status predicted <TM7>  
 F:983-1027/Domain: extracellular #status predicted <EXT>  
 F:379/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:512/Binding site: ATP (Lys) #status predicted  
 F:721,725,730/Active site: Asp, Asp, Lys #status predicted

Query Match 39.4%; Score 43; DB 1; Length 1027;  
 Best Local Similarity 60.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLRWISFC 16  
 DB 87 PTLRWKFC 96

## RESULT 41

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - fruit fly (*Drosophila melanogaster*)  
 N:Alternate names: sodium pump alpha chain  
 C:Species: *Drosophila melanogaster*  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
 C:Accession: S03632; S07049  
 R:Lebovitz, R.M.; Takeyasu, K.; Fambrough, D.M.  
 EMBO J. 8, 193-202, 1989

A:Title: Molecular characterization and expression of the (Na+K)-ATPase alpha-subunit in  
 A:Reference number: S03632; MUID:89231618; PMID:2540956  
 A:Accession: S03632  
 A:Molecule type: mRNA  
 A:Residues: 1-1038 <LEB>

A:Cross-references: UNIPROT:P13607; EMBL:X14476  
 A:Note: the sequence from Fig. 9 is inconsistent with that from Fig. 8 in having 89-Asp  
 A:Note: It is uncertain whether Met-1 or Met-40 is the initiator  
 R:Varad, A.; Gilmore-Heber, M.; Benz Jr., E.J.  
 FEBS Lett. 256, 203-207, 1989

A:Title: Amplification of the phosphorylation site - ATP-binding site cDNA fragment of  
 A:Reference number: S07049; MUID:9092469; PMID:2557235  
 A:Accession: S07049  
 A:Molecule type: mRNA  
 A:Residues: 397-521 <VAR>

A:Cross-references: EMBL:X17471  
 A:Note: the authors translated the codon ACC for residue 3 as Asn and AAT for residue 8  
 C:Genetics:

A:Gene: FlyBase:Atp-alpha  
 A:Cross-references: FlyBase:FBgn0002921  
 A:Map position: 3R 93B

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
 F:113-135/Domain: transmembrane #status predicted <TM1>  
 F:146-165/Domain: transmembrane #status predicted <TM2>  
 F:166-305/Domain: intracellular #status predicted <INT2>  
 F:306-328/Domain: transmembrane #status predicted <TM3>  
 F:335-363/Domain: transmembrane #status predicted <TM4>  
 F:364-801/Domain: intracellular #status predicted <INT3>  
 F:602-798/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:802-825/Domain: transmembrane #status predicted <TM5>  
 F:864-889/Domain: transmembrane #status predicted <TM6>  
 F:890-966/Domain: intracellular #status predicted <INT4>  
 F:967-993/Domain: transmembrane #status predicted <TM7>  
 F:994-1038/Domain: extracellular #status predicted <EXT>  
 F:391/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:732/Binding site: ATP (Lys) #status predicted  
 F:732,736,741/Active site: Asp, Asp, Lys #status predicted

Query Match 39.4%; Score 43; DB 1; Length 1038;  
 Best Local Similarity 44.4%; Pred. No. 2.4e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 5 DGPRLR-----EWISFC 16  
 DB 93 DGPRLRTPKOTPEWVKFC 110

## RESULT 42

hypochemical protein F42G8.6 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T32638  
 R:Gatting, S.; Holmes, A.

A:Description: The sequence of *C. elegans* cosmid F42G8.  
 A:Reference number: 221203  
 A:Accession: T32638

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-353 <GAT>  
 A:Cross-references: UNIPROT:O44510; EMBL:AF038618; PIDN:AA92067.1; GSPDB:GN00022; CESP  
 A:Experimental source: strain Bristol N2; clone F42G8  
 C:Genetics:

A:Gene: CESP:F42G8.6  
 A:Map position: 4  
 A:Insertion: 31/1; 71/3; 102/2; 225/1; 279/3  
 C:Superfamily: Porphyrin purpurea hypochemical protein 382

Query Match 39.0%; Score 42.5; DB 2; Length 353;  
 Best Local Similarity 38.9%; Pred. No. 1e+02;  
 Matches 7; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 3 CADGPTLR---EWISFC 17  
 DB 256 CGDNPTITAPIDVYLCFC 273

## RESULT 43

JH0470  
Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha chain (clone pAATNa136) - brine shrimp  
C:/Species: Artemia franciscana (brine shrimp)  
C:/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:/Accession: JH0470; S24196  
R:/Macias, M.T.; Palmero, I.; Sastre, L.  
Gene 105, 197-204, 1991  
A:/Title: Cloning of a cDNA encoding an Artemia franciscana Na/K ATPase alpha-subunit.  
A:/Reference number: JH0470; PMID:92039032; PMID:1657719  
A:/Accession: JH0470  
A:/Molecule type: mRNA  
A:/Residues: 1-1004 <MAC>  
A:/Cross-references: UNIPROT:P28774; EMBL:X56650; NID:910933; PIDN:CAA39972.1; PID:910934  
C:/Superfamily: Na+/K+-transporing ATPase alpha chain; ATPase nucleotide-binding domain  
C:/Keyword: ATP, heterodimer; hydrolyase; ion transport; phosphoprotein; potassium transp  
F:/2-1004/Product: Intracellular #status predicted <INT1>  
F:/2-75/Domain: Intracellular #status predicted <INT1>  
F:/76-97/Domain: transmembrane #status predicted <TM1>  
F:/111-130/Domain: transmembrane #status predicted <TM2>  
F:/321-271/Domain: transmembrane #status predicted <TM3>  
F:/372-296/Domain: transmembrane #status predicted <TM4>  
F:/301-329/Domain: transmembrane #status predicted <TM5>  
F:/330-767/Domain: Intracellular #status predicted <INT>  
F:/568-764/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:/768-791/Domain: transmembrane #status predicted <TM6>  
F:/830-855/Domain: transmembrane #status predicted <TM6>  
F:/856-936/Domain: intracellular #status predicted <INT7>  
F:/937-955/Domain: transmembrane #status predicted <TM7>  
F:/956-1004/Domain: extracellular #status predicted <EXT>  
F:/357/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:/489/Binding site: ATP (lys) #status predicted  
F:/598,702,707/Active site: Asp, Asp, Lys #status predicted

Query Match 39.0%; Score 42.5; DB 2; Length 1004;  
Best Local Similarity 47.4%; Pred. No. 2.7e+02;  
Matches 9; Conservative 0; Mismatches 3; Indels 7; Gaps 1;  
QY 5 DGP-----TLREWISFC 16  
DB 55 DGNCLTPKPTPEWIKFC 73

## RESULT 44

T00038  
hypothetical protein KIAA0289 - human (fragment)  
C:/Species: Homo sapiens (man)  
C:/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 05-Nov-1999  
C:/Accession: T00038  
R:/Ohara, O.; Nagase, T.; Ishikawa, K.; Nakajima, D.; Ohira, M.; Seki, N.; Nomura, N.  
submitted to the EMBL Data Library, August 1997  
A:/Description: Prediction of the coding sequences of unidentified human genes.  
A:/Reference number: Z14073  
A:/Accession: T00038  
A:/Status: preliminary; translated from GB/EMBL/DBJ  
A:/Molecule type: mRNA  
A:/Residues: 1-1302 <OHA>  
A:/Cross-references: EMBL:AB006627; NID:d1170680; PIDN:BA422958.1; PID:d1023834  
A:/Experimental source: Brain  
C:/Genetics:  
A:/Note: KIAA0289

Query Match 39.0%; Score 42.5; DB 2; Length 1302;  
Best Local Similarity 35.7%; Pred. No. 3.5e+02;  
Matches 10; Conservative 1; Mismatches 6; Indels 11; Gaps 1;  
QY 1 GGC-----ADGPTLREWISFCG 17  
DB 666 GGCCEQLCLQOMAPFPDDPTLYNIIIMFCG 693

## RESULT 45

## AH2829

conserved hypothetical protein Atu2063 [imported] - Agrobacterium tumefaciens (strain C  
C:/Species: Agrobacterium tumefaciens  
C:/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:/Accession: AH2829  
R:/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J  
Eriege, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:/Reference number: AB2577; PMID:21608550; PMID:11743193  
A:/Accession: AH2829  
A:/Status: preliminary  
A:/Molecule type: DNA  
A:/Residues: 1-141 <KUR>  
A:/Cross-references: UNIPROT:O8UD08; GB:AE008688; PIDN:ALL43054.1; PID:gl7740521; GSPDB:(  
A:/Experimental source: strain C58 (Dupont)  
C:/Genetics:  
A:/Gene: Atu2063  
A:/Map position: circular chromosome

Query Match 38.5%; Score 42; DB 2; Length 141;  
Best Local Similarity 47.1%; Pred. No. 53;  
Matches 8; Conservative 2; Mismatches 3; Indels 4; Gaps 1;  
QY 2 GCADGPTLREWISFCG 18  
DB 60 GAADAP----WLAFLCG 72

Search completed: September 1, 2005, 16:22:47  
Job time : 17.7266 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 66.9496 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-6

Perfect score: 109

Sequence: 1 GGCAADGPTLRKRWISFCGG 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Uniprot 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_tramb1:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	51.4	297	2 Q7UG64	Q7UG64 rhodopirell
2	54.5	50.0	934	2 Q9NE6	Q9NE6 caenorhabdi
3	50.5	46.3	387	2 Q9BA97	Q9BA97 rhizobium 1
4	50.5	46.3	389	2 O8KJF9	O8KJF9 rhizobium 1
5	50	45.9	386	1 ETR1_CANTR	O8Wzm candida tro
6	50	45.9	386	1 ETR2_CANTR	O8Wzm candida tro
7	49.5	45.4	405	2 Q9KIE9	Q9KIE9 streptomyc
8	49	45.0	245	2 Q9M060	Q9M060 arabidopsis
9	49	45.0	349	2 Q7V2B2	Q7V2B2 prochloroco
10	48	44.0	319	2 Q9RK45	Q9RK45 streptomyc
11	48	44.0	342	2 O6VME4	O6VME4 streptomyc
12	48	44.0	461	2 O7J2W7	O7J2W7 mycobacteri
13	48	44.0	1123	2 Q7QCC3	Q7QCC3 atropheles g
14	47.5	43.6	283	2 Q7ULR5	Q7ULR5 rhodopirell
15	47.5	43.6	283	2 O82CW2	O82CW2 streptomyc
16	47	43.1	94	2 O6MX73	O6MX73 azoarcus sp
17	47	43.1	129	2 O8DHX7	O8DHX7 synecococc
18	47	43.1	271	2 O89PR8	O89PR8 bradyrhizob
19	47	43.1	475	2 Q9UAT5	Q9UAT5 caenorhabdi
20	47	43.1	821	2 Q96D04	Q96D04 caenorhabdi
21	47	43.1	956	2 O6CLO9	O6CLO9 kluyveromyc
22	47	43.1	1926	2 Q9Y8B3	Q9Y8B3 paracoccidi
23	46.5	42.7	166	2 O6K939	O6K939 bacterioph
24	46.5	42.7	425	2 O89HD8	O89HD8 bradyrhizob
25	46	42.2	97	2 O8FPC4	O8FPC4 corynebacte
26	46	42.2	117	2 Q7MVA9	Q7MVA9 porphyromon
27	46	42.2	159	2 O8N852	O8N852 homo sapien
28	46	42.2	162	2 O63KH8	O63KH8 burkholderi
29	46	42.2	196	2 Q7VWWS	Q7VWWS bordetella
30	46	42.2	196	2 Q7W9K1	Q7W9K1 bordetella
31	46	42.2	245	2 O8GVFS	O8GVFS oryza sativ

32	46	42.2	275	2 O13090	O13090 pleurodeles
33	46	42.2	347	2 Q7PPP6	Q7PPP6 anopheles g
34	46	42.2	403	2 O88NU2	O88NU2 pseudomonas
35	46	42.2	443	2 Q9P858	Q9P858 phaeosphaer
36	46	42.2	482	2 Q6AIT0	Q6AIT0 desulfotale
37	46	42.2	926	1 AASS_HUMAN	AASS_HUMAN
38	46	42.2	1902	2 Q9Y878	Q9Y878 coccidioid
39	45.5	41.7	309	2 O8XZNS	O8XZNS ralsnola s
40	45	41.3	108	2 Q7RUN5	Q7RUN5 neurospora
41	45	41.3	16	2 Q6ZTT4	Q6ZTT4 homo sapien
42	45	41.3	173	2 O8C4M6	O8C4M6 mus musculu
43	45	41.3	180	2 Q07291	Q07291 natronomona
44	45	41.3	209	2 O6N1X5	O6N1X5 rhodopseudo
45	45	41.3	290	2 O89JRS	O89JRS bradyrhizob
46	45	41.3	338	2 O82CX1	O82CX1 streptomyc
47	45	41.3	379	2 Q7SKV0	Q7SKV0 brachydanio
48	45	41.3	410	2 Q6Z9V1	Q6Z9V1 burkholderi
49	45	41.3	410	2 Q63H26	Q63H26 burkholderi
50	45	41.3	421	2 Q9XUV7	Q9XUV7 caenorhabdi
51	45	41.3	499	1 MEP2_YEAST	MEP2_YEAST
52	45	41.3	540	2 O82LTO	O82LTO streptomyc
53	45	41.3	594	2 Q7SHC4	Q7SHC4 neurospora
54	45	41.3	769	2 Q70804	Q70804 et virus. 1
55	44.5	40.8	175	2 Q7XQ02	Q7XQ02 oryza sativ
56	44.5	40.8	248	2 Q7PXF4	Q7PXF4 anopheles g
57	44.5	40.8	282	2 Q7QCK2	Q7QCK2 cheinus the
58	44.5	40.8	429	2 Q8AVB0	Q8AVB0 brachydanio
59	44.5	40.8	485	2 O8SC10	O8SC10 propionibac
60	44.5	40.8	497	2 Q7QIK7	Q7QIK7 anopheles g
61	44.5	40.8	818	2 Q6PBA6	Q6PBA6 brachydanio
62	44.5	40.8	1067	2 O6MZ09	O6MZ09 aspergillus
63	44.5	40.8	1142	2 O8ITL2	O8ITL2 trypanosoma
64	44.5	40.8	1163	2 Q7NLE9	Q7NLE9 gloeobacter
65	44.5	40.8	173	2 Q6ZAD7	Q6ZAD7 oryza sativ
66	44	40.4	173	2 Q6QHD2	Q6QHD2 gallid herp
67	44	40.4	178	2 Q6PL14	Q6PL14 gallid herp
68	44	40.4	197	2 Q6R8A0	Q6R8A0 sodalis glo
69	44	40.4	209	2 Q9L059	Q9L059 streptomyc
70	44	40.4	210	2 Q69PA9	Q69PA9 oryza sativ
71	44	40.4	238	2 Q7QKX0	Q7QKX0 anopheles g
72	44	40.4	277	1 IMP1_MOUSE	IMP1_MOUSE
73	44	40.4	292	2 Q67642	Q67642 gallid herp
74	44	40.4	298	2 Q66553	Q66553 gallid herp
75	44	40.4	310	2 Q678H2	Q678H2 lymphocyeti
76	44	40.4	346	2 Q62030	Q62030 caenorhabdi
77	44	40.4	371	2 Q9RRJ3	Q9RRJ3 deinococcus
78	44	40.4	385	2 Q7XMK0	Q7XMK0 oryza sativ
79	44	40.4	404	2 Q7QFA0	Q7QFA0 anopheles g
80	44	40.4	425	2 Q8PD03	Q8PD03 xanthomonas
81	44	40.4	450	2 Q75211	Q75211 ashbya goss
82	44	40.4	490	2 Q04270	Q04270 chlamydomon
83	44	40.4	519	2 Q7YIN9	Q7YIN9 oryza sativ
84	44	40.4	524	2 Q66GJ0	Q66GJ0 arabidopsis
85	44	40.4	524	2 Q84W33	Q84W33 arabidopsis
86	44	40.4	524	2 Q9C868	Q9C868 arabidopsis
87	44	40.4	537	2 Q63M37	Q63M37 burkholderi
88	44	40.4	613	2 Q7YT25	Q7YT25 dirosophila
89	44	40.4	613	2 Q9VGR8	Q9VGR8 dirosophila
90	44	40.4	742	2 Q8SRJ4	Q8SRJ4 encephalit
91	44	40.4	926	1 AASS_MOUSE	AASS_MOUSE
92	44	40.4	946	2 Q9UG55	Q9UG55 mus musculu
93	44	40.4	974	1 PHS2_SOLITU	PHS2_SOLITU
94	44	40.4	997	2 Q6BI26	Q6BI26 solanum tub
95	44	40.4	1008	2 Q8AY57	Q8AY57 fundulus he
96	44	40.4	1011	2 Q6VYV7	Q6VYV7 oncorhynch
97	44	40.4	1022	1 ATIA_TORCA	ATIA_TORCA
98	44	40.4	1023	1 ALAI_HUMAN	ALAI_HUMAN
99	44	40.4	1025	2 Q7ZYK8	Q7ZYK8 xenopus lae
100	44	40.4			

## ALIGNMENTS

```

RESULT 1
Q7U0E4 PRELIMINARY; PRT; 297 AA.
AC Q7U0E4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=RB6375;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxId=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294144; CAD74759.1; -.
DR InterPro; IPR00194; ATPase_a/bcentre.
DR PROSITE; PS00152; ATPase_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50829; GYF; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475F670F02C78B9B CRC64;

Query Match 51.4%; Score 56; DB 2; Length 297;
Best local Similarity 69.2%; Pred. No. 2.2;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCADGPTLRWIS 14
Db 173 GPADGPTMKQWIS 185

RESULT 2
Q9NEX6 PRELIMINARY; PRT; 934 AA.
AC Q9NEX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Hypothetical protein Y105E8A.21.
GN ORFNames=Y105E8A.21;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132876; CAC48140.1; -.
DR WormBase; WBGenome00013679; Y105E8A.21.
DR Wormpep; Y105E8A.21; C225162.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; C:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.

```

```

KW Hypothetical protein.
SQ SEQUENCE 934 AA; 10485 MW; 5ED4E1D03DB06F24 CRC64;

Query Match 50.0%; Score 54.5; DB 2; Length 934;
Best local Similarity 58.8%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 3 CADGPTLRW-ISPCCG 18
Db 899 CVDGTSRDMPVPSFTCG 915

RESULT 3
Q98A97 PRELIMINARY; PRT; 387 AA.
AC Q98A97;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE M16096 protein.
GN OrderedlocusNames=m16096;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Mesorhizobium.
OX NCBI_TaxId=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Agamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iessawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakasaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003008; BAB52440.1; -.
DR HSSP; P77407; 1POY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CA1B BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
DR Complete proteome.
SQ SEQUENCE 387 AA; 42226 MW; 64643EBBC8F25518 CRC64;

Query Match 46.3%; Score 50.5; DB 2; Length 387;
Best local Similarity 42.9%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

QY 3 CADGPTLR-----REMISFC 16
Db 237 CADGKEIVFSVQNDREVMVNC 257

RESULT 4
Q8KJF9 PRELIMINARY; PRT; 389 AA.
AC Q8KJF9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE PUTATIVE RACEMASE/DEHYDRATASE PROTEIN.
GN Name=ms1181;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Mesorhizobium.
OX NCBI_TaxId=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R7A;
RX MEDLINE=21999272; PubMed=12003951;
RX DOI=10.1128/JB.184.11.3086-3095.2002;
RA Sullivan J.T., Trzciatkowski J.R., Crutickbank R.W., Gouzy J.,

```

RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rosebach U.,  
 RA Stuart G.S., Weaver J.B., Webby R.J., de Bruijn F.J., Ronson C.W.;  
 RT "Comparative sequence analysis of the symbiosis island of  
 RT Mesorhizobium loti strain R7A."  
 RL J. Bacteriol. 184:3086-3095(2002).  
 DR EMBL: AL672113; CAD31586.1; -  
 DR HSSP: P77407.1PQY.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR InterPro: IPR003673; CA1B\_BAIF.  
 DR Pfam: PF02515; COA\_transf\_3; 1.  
 SQ SEQUENCE 389 AA; 42703 MW; 6678D2C96A7E5204 CRC64;  
 Query Match 46.3%; Score 50.5; DB 2; Length 389;  
 Best Local Similarity 42.9%; Pred. No. 21;  
 Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;  
 Oy 3 CADGPTL-----REMISFC 16  
 |||||:|||||  
 Db 243 CADGKEVIFSVQNDREWVNF 263  
 RESULT 5  
 ETR1\_CANTR STANDARD; PRT; 386 AA.  
 AC Q8WZM3;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Enoyl-[acyl-carrier protein] reductase [NADPH, B-specific] 1,  
 DE mitochondrial precursor (EC 1.3.1.10).  
 GN Name=ETR1;  
 OS Candida tropicalis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5482;  
 [1]  
 RN SEQUENCE FROM N.A., SEQUENCE OF 23-29, FUNCTION, SUBUNIT, AND  
 RP SUBCELLULAR LOCATION.  
 RC STRAIN=ATCC 20336;  
 RX PubMed=12890667; DOI=10.1074/jbc.M307664200;  
 RA Torokko J.M., Kojivuranta K.T., Kaestliotis A.J., Airene T.T.,  
 RA Glumoff T., Ilves M., Hartig A., Gurvitz A., Hiltunen J.K.;  
 RT "Candida tropicalis expresses two mitochondrial 2-enoil thioester  
 RT reductases that are able to form both homodimers and heterodimers";  
 RL J. Biol. Chem. 278:41213-41220(2003).  
 [2]  
 RN SUBUNIT.  
 RP STRAIN=ATCC 20336;  
 RX PubMed=12890667; DOI=10.1074/jbc.M307664200;  
 RA Airene T.T., Torokko J.M., Van den Plas S., Sotomunen R.T.,  
 RA Kaestliotis A.J., Wierenga R.K., Hiltunen J.K.;  
 RT "Structure-function analysis of enoyl thioester reductase involved in  
 RT mitochondrial maintenance";  
 RL J. Mol. Biol. 327:47-59(2003).  
 [3]  
 RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS), AND MUTAGENESIS OF TYR-79.  
 RP PubMed=12614607; DOI=10.1016/S0022-2836(03)00038-X;  
 RA Airene T.T., Torokko J.M., Van den Plas S., Sotomunen R.T.,  
 RA Kaestliotis A.J., Wierenga R.K., Hiltunen J.K.;  
 RT "Structure-function analysis of enoyl thioester reductase involved in  
 RT mitochondrial maintenance";  
 RL J. Mol. Biol. 327:47-59(2003).  
 [4]  
 CC -1- FUNCTION: Required for respiration and the maintenance of the  
 CC mitochondrial compartment. May have a role in the mitochondrial  
 CC synthesis of fatty acids.  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.  
 CC -1- SUBUNIT: Homodimer and heterodimer with etz2.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrion.  
 CC -1- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase  
 CC family. Quinone oxidoreductase subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL: U94997; AAL55472.1; -  
 DR PDB: 1GU7; X-ray; A/B=23-386.  
 DR PDB: 1GUP; X-ray; A/B=23-386.  
 DR PDB: 1GYR; X-ray; A/B/C=23-386.  
 DR InterPro: IPR002085; Adh\_zn\_family.  
 DR InterPro: IPR010324; GroES\_Like.  
 DR Pfam: PF00107; ADH\_zinc\_N; 1.  
 KW 3D-structure; Direct protein sequencing; Fatty acid biosynthesis;  
 KW Mitochondrion; NADP; Oxidoreductase; Transit peptide.  
 FT TRANSIT 1 22  
 FT CHAIN 23 386  
 FT MUTAGEN 79 79  
 FT SEQUENCE 386 AA; 42160 MW; FCBC174A240742D8 CRC64;  
 SQ  
 Query Match 45.9%; Score 50; DB 1; Length 386;  
 Best Local Similarity 61.5%; Pred. No. 25;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 6 GPTLEWISPCGG 18  
 |||||:|||||  
 Db 254 GPTLEWIKOSGG 266  
 RESULT 6  
 ETR2\_CANTR STANDARD; PRT; 386 AA.  
 AC Q8WZM4;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Enoyl-[acyl-carrier protein] reductase [NADPH, B-specific] 2,  
 DE mitochondrial precursor (EC 1.3.1.10).  
 GN Name=ETR2;  
 OS Candida tropicalis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5482;  
 [1]  
 RN SEQUENCE FROM N.A., FUNCTION, AND SUBUNIT.  
 RP STRAIN=ATCC 20336;  
 RX PubMed=12890667; DOI=10.1074/jbc.M307664200;  
 RA Torokko J.M., Kojivuranta K.T., Kaestliotis A.J., Airene T.T.,  
 RA Glumoff T., Ilves M., Hartig A., Gurvitz A., Hiltunen J.K.;  
 RT "Candida tropicalis expresses two mitochondrial 2-enoil thioester  
 RT reductases that are able to form both homodimers and heterodimers";  
 RL J. Biol. Chem. 278:41213-41220(2003).  
 [2]  
 RN X-RAY CRYSTALLOGRAPHY (2.11 ANGSTROMS).  
 RP Airene T.T., Torokko J.M., Hiltunen J.K.;  
 RT "Crystal structure of enoyl thioester reductase 2.";  
 RL Submitted (JUN-2002) to the PDB data bank.  
 [3]  
 CC -1- FUNCTION: Required for respiration and the maintenance of the  
 CC mitochondrial compartment. May have a role in the mitochondrial  
 CC synthesis of fatty acids.  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.  
 CC -1- SUBUNIT: Homodimer and heterodimer with ETR1.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrion (By similarity).  
 CC -1- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase  
 CC family. Quinone oxidoreductase subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U94996; AAL55471.1; -.

DR PDB; 1H0K; X-ray; A/B=23-386.

DR InterPro; IPR002085; Adh zn family.

DR InterPro; IPR011032; GroES-like.

DR Pfam; PF00107; ADH zinc N; 1.

KM 3D-structure; fatty acid biosynthesis; Mitochondrion; NADP; Oxidoreductase; Transit peptide.

FT TRANSIT 1 22 Mitochondrion (potential)

FT CHAIN 23 386 Enoyl-[acyl-carrier protein] reductase [NADPH, B-specific] 2.

FT

SEQ SEQUENCE 386 AA; 42116 MW; 91ABE00831F0CE8 CRC64;

Query Match 45.9%; Score 50; DB 1; Length 386;  
Best Local Similarity 61.5%; Pred. No. 25;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GPTLRWISFCG 18  
DB 254 GPTKEIKQSG 266

RESULT 7

ID Q9KIE9 PRELIMINARY; PRT; 405 AA.

AC Q9KIE9;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE PKbE.

GN Name=fkBE;

OS Streptomyces hygroscopicus subsp. ascomyceticus.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycinae; Streptomycetaceae; Streptomycetes.

OX NCBI\_TaxId=132248;

OX (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=20323220; PubMed=10863099; DOI=10.1016/S0378-1119(00)00171-2;

RT "The FK520 gene cluster of Streptomyces hygroscopicus var. Wu K., Chung L., Revill W.P., Katz L., Reeves C.D."

RT "The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units."

RL Gene 251;81-90(2000).

RL EMBL; AF235504; AAF86384.1; -.

DR HSSP; P77407; 1POY.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR003673; CAIB BAIF.

DR Pfam; PF02515; CoA\_transf\_3; 1.

SEQ SEQUENCE 405 AA; 43696 MW; DC2569DFC914AD6F CRC64;

Query Match 45.4%; Score 49.5; DB 2; Length 405;  
Best Local Similarity 50.0%; Pred. No. 31;  
Matches 10; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

QY 5 DGPTL-----REWISFCG 17  
DB 252 DGQTNLGLONEREMASFCG 271

RESULT 8

ID Q9M060 PRELIMINARY; PRT; 245 AA.

AC Q9M060;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Eukaryotic translation initiation factor 6 (EIF-6)-like protein (A1355620).

GN Name=Flit6\_30; Synonyms=At3g55620;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxId=3702;

RN (1)

RP SEQUENCE FROM N.A.

RA Benes V., Wurbach E., Drzonek H., Ansoerge W., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN (2)

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN (3)

RP SEQUENCE FROM N.A.

RA Shin P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P., Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

RN (4)

RP SEQUENCE FROM N.A.

RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M., Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H., Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamita A., Kawai J., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K., Ecker J., Theologis A., Davis R.W.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL161667; CAB81587.1; -.

DR EMBL; BT009656; AAP75806.1; -.

DR EMBL; AY128351; AAM91554.1; -.

DR PIR; T47701; T47701.

DR HSSP; Q12522; 1G62.

DR GO; GO:0003743; P:translation initiation factor activity; IEA.

DR GO; GO:0006413; P:translational initiation; IEA.

DR InterPro; IPR002769; eIF6.

DR Pfam; PF01912; eIF-6; 1.

DR ProDom; PD006880; eIF6; 1.

DR SMART; SM00654; eIF6; 1.

DR TIGRPFAM; TIGR00323; eIF-6; 1.

KM Initiation factor.

SEQ SEQUENCE 245 AA; 26482 MW; 73369A2A657F390D CRC64;

Query Match 45.0%; Score 49; DB 2; Length 245;  
Best Local Similarity 57.1%; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 ADGPTLRWISFCG 17  
DB 194 AAGMTVNDWTSFCG 207

RESULT 9

ID Q7V2B2 PRELIMINARY; PRT; 349 AA.

AC Q7V2B2;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Dihydroorotase (EC 3.5.2.3).

GN Name=pyrC; Ordered locus names=pmw0569;

OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).

OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcales;

OC Prochlorococcus.

OX NCBI\_TaxId=59919;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=2825658; PubMed=12917642; DOI=10.1038/nature01947;

RA Raccap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,  
 RA Ailgren N.A., Ariellano A., Coleman M., Hauser L., Hees W.R.,  
 RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,  
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
 RA Webb E.A., Zinner B.R., Chisholm S.W.;  
 RT "genome divergence in two *Prochlorococcus* ecotypes reflects oceanic  
 niche differentiation.";   
 RT Nature 424:1042-1047(2003).  
 RL EMBL; BX572091; CAS19028.1; -.  
 DR HSSP; P05020; 1J79.  
 DR GO; GO:0004151; F:diaphorase activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0019856; F:pyrimidine base biosynthesis; IEA.  
 DR InterPro; IPR006680; Amidohydro\_1.  
 DR InterPro; IPR004721; DHD1mr.  
 DR Pfam; PF01979; Amidohydro\_1; 1.  
 DR Trifam; TRF00856; pyrc\_dimer; 1.  
 DR PROSITE; PS00482; DIHYDROOROTASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00483; DIHYDROOROTASE\_2; 1.  
 KM Complete proteome; Hydrolase.  
 SQ SEQUENCE 349 AA; 39958 MW; CC02P5AAE02EC927 CRC64;  
 \*  
 Query Match 45.0%; Score 49; DB 2; Length 349;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 \*  
 Qy 2 GCADGPTLRWISFC 17  
 Db 243 GTDSAPHLRQMKAFGC 258  
 \*  
 RESULT 10  
 Q9RKM5 PRELIMINARY; PRT; 319 AA.  
 AC Q9RKM5; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Putative Merf family transcriptional regulator.  
 GN OPRNames=SCD17.06c;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteriae; Actinobacteriales; Streptomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares R., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete *Streptomyces*  
 RT *coelicolor* A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- SIMILARITY: Contains 1 HTH merf-type DNA-binding domain.  
 DR EMBL; AL939118; CAB56383.1; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003760; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000551; HTH\_Merf.  
 DR Pfam; PF00376; Merf; 1.  
 DR PRINTS; PR00040; HTMERF.  
 DR SMART; SM00422; HTH\_MERF; 1.  
 DR PROSITE; PS50937; HTH\_MERF\_2; 1.  
 KM Complete proteome; DNA-binding.

SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;  
 \*  
 Query Match 44.0%; Score 48; DB 2; Length 319;  
 Best Local Similarity 61.5%; Pred. No. 42;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 \*  
 Qy 2 GCADGPTLRWIS 14  
 Db 255 GRDGPRLRWELA 267  
 \*  
 RESULT 11  
 Q6VNH4 PRELIMINARY; PRT; 342 AA.  
 AC Q6VNH4; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Putative SARF family pathway specific regulatory protein.  
 GN Name=slp;  
 OS Streptomyces ambofaciens.  
 OC Bacteria; Actinobacteriae; Actinobacteriales; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=1889;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 23877;  
 RX PubMed=14742212;  
 RA Pang X., Aigle B., Girardet J.M., Mangenot S., Pernodet J.L.,  
 RA Decaris B., Lebiond P.;  
 RT "Functional angucycline-like antibiotic gene cluster in the terminal  
 RT inverted repeats of the *Streptomyces ambofaciens* linear chromosome.";  
 RL Antimicrob. Agents Chemother. 48:575-588(2004).  
 DR EMBL; AY338477; AAR30165.1; -.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
 DR InterPro; IPR009059; bi\_resp\_regltr\_C.  
 DR InterPro; IPR00158; BTAD.  
 DR Pfam; PF03704; BTAD; 1.  
 DR Pfam; PF00486; Trans\_reg\_C; 1.  
 SQ SEQUENCE 342 AA; 35635 MW; 945BC929E5A6E3D CRC64;  
 \*  
 Query Match 44.0%; Score 48; DB 2; Length 342;  
 Best Local Similarity 46.7%; Pred. No. 45;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 \*  
 Qy 2 GCADGPTLRWISFC 16  
 Db 112 GCGGPGSPRPLWESC 126  
 \*  
 RESULT 12  
 Q73ZW7 PRELIMINARY; PRT; 461 AA.  
 AC Q73ZW7; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=MAP1484c;  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteriae; Actinobacteriales; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=k10;  
 RA Li L., Baumann J., Zhang Q., Amosin A., Alt D., Kapur V.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE017232; AAS03801.1; -.  
 DR GO; GO:0005506; P:iron ion binding; IEA.

DR GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO:0006725; P:aromatic compound metabolism; IEA.  
 DR GO:0006118; P:electron transport; IEA.  
 DR InterPro: IPR005806; Rieseke, reg.  
 DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
 DR Pfam: PF00355; Rieseke; 1.  
 DR PRINTS: PR00090; RINGDIKXGNASE.  
 DR Complete proteome.  
 KW  
 SQ SEQUENCE 461 AA; 52010 MW; 208E39A89C121839 CRC64;

Query Match 44.0%; Score 48; DB 2; Length 461;  
 Best Local Similarity 47.4%; Pred. No. 60;  
 Matches 9; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 1 GGCAGWLNDDADPALDMM 174  
 156 GGCAGWLNDDADPALDMM 174

## RESULT 13

Q70C63 PRELIMINARY; PRT; 1123 AA.  
 AC Q70C63;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE AgCP1221.  
 GN Name=agCG53078; ORFNames=ENSANG0000018866;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST.  
 RA Anopheles Genome Sequencing Consortium;  
 RB Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.

DR EMBL: AAB0100859; EAA08177.1; -.  
 DR GO:0005524; F:ATP binding; IEA.  
 DR GO:0008898; F:homocysteine S-methyltransferase activity; IEA.  
 DR GO:0004672; F:protein kinase activity; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR003726; S\_methyl\_trans.  
 DR Pfam: PF00069; Kinase; 1.  
 DR Pfam: PF02574; S-methyl\_trans; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 SQ SEQUENCE 1123 AA; 12006 MW; D3CC001D8D4882AF CRC64;

Query Match 44.0%; Score 48; DB 2; Length 1123;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ADGPTLRWTISF 15  
 969 ADGPTLRWTISF 980

## RESULT 14

Q7ULR5 PRELIMINARY; PRT; 238 AA.  
 AC Q7ULR5;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Similar to phycoerythrin alpha phycoerythrin lyase CpeC (EC 4.-.-.-  
 DE ).  
 GN Name=cpeC; OrderedLocustNames=RB9340;

OS Rhodospirillum rubrum.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).

DR EMBL: BX94149; CAD76204.1; -.  
 DR GO:0016829; F:lyase activity; IEA.  
 DR InterPro: IPR008938; ARM.  
 DR InterPro: IPR004155; PBS\_lyase\_HEAT.  
 DR Pfam: PF01330; HEAT\_PBS; 1.  
 DR SMART: SM00567; EZ\_HEAT; 3.  
 DR Complete proteome; Lyase.  
 KW  
 SQ SEQUENCE 238 AA; 26142 MW; B7CA7284593B0C72 CRC64;

Query Match 43.6%; Score 47.5; DB 2; Length 238;  
 Best Local Similarity 39.1%; Pred. No. 37;  
 Matches 9; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

QY 1 GGCADGP-----TLREM 12  
 29 GGCADGPVYALKHNPFTMRQW 51

## RESULT 15

Q82CW2 PRELIMINARY; PRT; 283 AA.  
 AC Q82CW2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative ICLR-family transcriptional regulator.  
 DE OrderedLocustNames=SAV5226;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433196;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.,  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RT Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL: AP005042; BAC72938.1; -.  
 DR GO:0003677; F:DNA binding; IEA.  
 DR GO:0006555; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR005471; HTH\_ICIR.  
 DR InterPro: IPR009058; Wing\_hlx\_DNA\_bnd.  
 DR Pfam: PF01614; ICIR; 1.  
 DR Complete proteome.  
 KW  
 SQ SEQUENCE 283 AA; 30503 MW; F63B1705578EE67 CRC64;

Query Match	Similarity	Score	DB	Length
Best Local	50.0%	47.5	2	283
Matches	8; Conservative	Pred. No. 44;		
		3; Mismatches	2;	Indels 3; Gaps 1;
QY	3 CADGPT---LRWISF	15		
	:    :    :			
Db	152 CAGGPTPAVHWVDF	167		

## RESULT 16

	PRELIMINARY:	PRT:	94 AA.
Q6MX73;			
AC Q6MX73;			
05-JUL-2004 (TREMBLrel. 27, Created)			
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DE Hypothetical protein.			
GN ORFNames=C2B002;			
OS Azococcus sp. (strain EbN1).			
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;			
OC Rhodocyclaceae; Azococcus.			
OX NCBI_TaxID=76114;			
OX [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=EbN1;			
RA Kube M., Heider J., Amann J., Hufnagel P., Kuehner S., Beck A.			
RA Reinhardt R., Rabus R.;			
RL submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.			
RL [2]			
RP SEQUENCE FROM N.A.			
Rt STRAIN=EbN1;			
RA PROSCIENCE;			
RL submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BX682953; CAF21985.1; -.			
KW Hypothetical protein.			
SQ SEQUENCE 94 AA; 9829 MW; 91AC4ECCAB8A7FDE CRC64;			
Query Match	43.1%;	Score 47;	DB 2; Length 94;
Best Local Similarity	72.7%;	Pred. No. 18;	
Matches 8	Conservative 1;	Mismatches 2;	Indels 0;

## RESULT 17

ID	Q8DHX7	PRELIMINARY;	PR;	129	AA.
AC	Q8DHX7				
DT	01-MAR-2003	(TREMBLrel. 23, Created)			
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	T11816 protein.				
CN	OrderedLocusName=t11816;				
OS	Synechococcus elongatus (Thermosynechococcus elongatus).				
OC	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.				
OX	NCBI_TaxID=32046;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BP-1;				
RX	MEDLINE=22225144; PubMed=12440834;				
RA	Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.				
RA	Kiyokawa A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,				
RA	Matsumoto M., Matsuno M., Matsuno A., Nakazaki N.,				
RA	Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;				
RT	"Complete genome structure of the thermophilic cyanobacterium				
RT	Thermosynechococcus elongatus BP-1."				
RL	DNA Res. 9:123-130(2002)				
KR	EMBL, AP005375; BAC09358.1; -.				
DR	Complete proteome.				
SQ	SEQUENCE 129 AA; 14644 MW; EBB44691E7DD1E12 CRC64;				

RESULT 18

ID	089PE8:	PRELIMINARY;	PRF,	271 AA.
DT	01-JUN-2003 (TEMBLrel. 24, Created)			
DT	01-JUN-2003 (TEMBLrel. 24, Last sequence update)			
DT	01-OCT-2003 (TEMBLrel. 25, Last annotation update)			
DE	CutM protein.			
GN	Name=cutM; OrderedLocusNames=bl1r3534;			
OS	Bradyrhizobium japonicum.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bradyrhizobiaceae; Bradyrhizobium.			
OX	NCBI_TaxID=375;			
RY	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=USDA110.			
RX	MEDLINE=24484998; PubMed=12597275;			
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiimi T.,			
RA	Sasanoto S., Metanabe A., Ideawa K., Iriyuchi M., Kawashima K.,			
RA	Kohata M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada			
RA	Takaba S.;			
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium			
RT	Bradyrhizobium japonicum USDA110."			
RL	DNA Res. 9:189-197 (2002).			
DR	EMBL; AP005548; BAC48799.1; -			
DR	HSP; F19920; IN62.			
DR	GO; GO:0016491; P:oxidoreductase activity; IEA.			
DR	GO; GO:0006118; P:electron transport; IEA.			
DR	Interpro; IPR005107; CO deh flav C.			
DR	Interpro; IPR002346; CO dehydrog. molyb.			
DR	Pfam; PF00941; FAD_binding_5; 1.			
DR	Complete proteome.			
QO	SEQUENCE 271 AA; 29422 MW; 4959C9F9A814FDC6 CRC64;			

RESULT 19  
CONVATE

ID	09UAT5	PRELIMINARY;	PRT;	475 AA.
AC	09UAT5			
DT	01-MAY-2000	(TREMBLrel_13, Created)		
DT	01-MAY-2000	(TREMBLrel_13, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel_25, Last annotation update)		
DE	Hypothetical protein C01B4.7.			
GN	Name=C01B4.7; ORFNames=C01B4.7;			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metacasta; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderiinae; Caenorhabditis.			
OX	NCBI_TaxId=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Bristol N2;			
FX	MEDLINE=99068613; Pubmed=9851916;			
RG	WormBase Consortium;			
RT	"Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";			



```

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Smith A., Mameley P., Fronick W.;
RT "The sequence of C. elegans cosmid C01B4.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: ARI25952; AAD14699.1; -.
DR PIR: T33943; T33943.
DR WormBase; WBGene0015271; C01B4.7.
DR WormPep; C01B4.7; CE20476.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
DR Hypothetical protein.
KM SEQUENCE 475 AA; 53094 MW; 79095D45572AF535 CRC64;
SQ
Query Match 43.1%; Score 47; DB 2; Length 475;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 3 CADGPTLRWISFCGG 18
Db 268 CTDRCVLSAWVSFLGG 283
.
```

```

RA Waterston R.H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006722; AAK68417.1; -.
DR PDB; 1LUR; X-ray; A/B=483-821.
DR WormBase; WBGene0021219; Y19D10A.4.
DR WormPep; Y19D10A.4; CE21450.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004034; F:aldose 1-epimerase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR008183; Ald1 epimerase.
DR InterPro; IPR011013; Gal_mut1-like.
DR InterPro; IPR007114; MFS.
DR Pfam; PF01263; Aldose_epim; 1.
DR PROSITE; PS50850; MFS; 1.
DR Hypothetical protein.
KM SEQUENCE 821 AA; 91593 MW; 923A78BFC95D1A76 CRC64;
SQ
Query Match 43.1%; Score 47; DB 2; Length 821;
Best Local Similarity 50.0%; Pred. No. 1,5e+02;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 3 CADGPTLRWISFCGG 18
Db 268 CTDRCVLSAWVSFLGG 283
.
```

```

RESULT 21
Q6CLJ9 PRELIMINARY; PRT; 956 AA.
ID Q6CLJ9
AC Q6CLJ9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P40825 Saccharomyces cerevisiae YOR333cc ALA1 alanyl-tRNA
DE synthetase.
GN ORFNames=KLA0F024319;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
Latoune I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
```



RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boissiere A., Boyer J., Catellico L., Confarotier F., de Danovar A.,  
 RA Despres L., Fabre B., Fairhead C., Ferry-Dumazet H., Giropi A.,  
 RA Hantreay F., Henneguier C., Jaumaux N., Joyet P., Kachouri R.,  
 RA Kerret A., Kozan R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicard J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straud M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Soulet J.L.,  
 RL "genome evolution in yeasts.",  
 RL Nature 430:35-44 (2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RA Genoscope (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382126; CAG97897.1; -.  
 DR GO; GO:0004813; F:alanine-tRNA ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; F:amyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR003156; Pesterase\_DHHAL.  
 DR InterPro; IPR02318; tRNA-synt\_2c.  
 DR Pfam; PF02272; DHHA1; 1. tRNA\_synt\_Ala.  
 DR Pfam; PF01411; tRNA-synt\_2c; 1.  
 DR PRINTS; PR00860; TRNA-SYNTALA.  
 DR TIGRFAMs; TIGR00344; alas; 1.  
 DR PROSITE; PS50860; AA-tRNA\_LIGASE\_II\_ALA; 1.  
 KM Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 956 AA; 107100 MW; 4F5CE6585880A3C CRC64;

Query Match 43.1%; Score 47; DB 2; Length 956;  
 Best Local Similarity 50.0%; Pred. No. 1.Be+02;  
 Matches 9; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

OY 5 DGPUREW---ISFCG 18  
 : ||| |||  
 Db 704 ENPTSEKOKSIKIFPCG 721

RESULT 22  
 O9Y8B3 PRELIMINARY; PRT; 1926 AA.  
 ID O9Y8B3;  
 AC O9Y8B3;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Glucan synthase.  
 GN Name=Fks;  
 OS Paracoccidioides brasiliensis.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Omygenales; mitosporic Omygenales; Paracoccidioides.  
 CX NCBI\_Taxid=121759;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pb01;  
 RX MEDLINE=20171859; PubMed=10705373;  
 RX DOI=10.1002/SITC.1097-0061(200003)016:5<451::AID-YEAS40>3.0.CO;2-O;  
 RA Pereira M., Felipe M.S.S., Brigido M.M., Soares C.M.A., Azevedo M.O.;  
 RT "Molecular cloning and characterization of a glucan synthase gene from  
 the human pathogenic fungus Paracoccidioides brasiliensis.";  
 RL Yeast 16:451-462(2000).  
 DR EMBL; AF148715; AAD37783.1; -.  
 DR GO; GO:0000148; C:1,3-beta-glucan synthase complex; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0003843; F:1,3-beta-glucan synthase activity; IEA.  
 DR GO; GO:0006075; F:beta-1,3 glucan biosynthesis; IEA.  
 DR InterPro; IPR003440; Glyco\_trans\_48.  
 DR InterPro; IPR002114; Hpt\_Serp\_S.

DR Pfam; PF02364; Glucan synthase; 1.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
 SQ SEQUENCE 1926 AA; 220574 MW; BB098550FD2253D5 CRC64;

Query Match 43.1%; Score 47; DB 2; Length 1926;  
 Best Local Similarity 46.7%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 GCADGPTLRWISFC 16  
 : ||| : |||  
 Db 1374 GCADTPTIDRWVRC 1388

RESULT 23  
 O6KG99 PRELIMINARY; PRT; 166 AA.  
 ID O6KG99;  
 AC O6KG99;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Bacteriophage Felix 01.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.  
 CX NCBI\_Taxid=77775;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stranganathan N., Whitchard J.M., Pierson F.W., Kapur V., Weigt L.A.;  
 RT "Bacteriophage Felix 01: Genetic Characterization.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF320576; AAO14824.1; -.  
 KM Hypothetical protein.  
 SQ SEQUENCE 166 AA; 19296 MW; 5AAB33B39DC3C989 CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 166;  
 Best Local Similarity 52.9%; Pred. No. 37;  
 Matches 9; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

OY 1 GGCADGPTLRWISFC 17  
 : ||| |||  
 Db 8 GSC---PTYGHWISLCG 21

RESULT 24  
 O89HD8 PRELIMINARY; PRT; 426 AA.  
 ID O89HD8;  
 AC O89HD8;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE B1r6053 protein.  
 GN OrderedLocustNames=B1r6053;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Bradyrhizobiaceae; Bradyrhizobium.  
 CX NCBI\_Taxid=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDAL10;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasano S., Watanabe A., Ideasa K., Iriuguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimo S., Tsunoda H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 Bradyrhizobium japonicum USDAL10.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AP005957; BAC51318.1; -.  
 DR HSSP; P27017; 1000.  
 KM Complete proteome.  
 SQ SEQUENCE 426 AA; 47042 MW; AE20A1BC6CBE038 CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 426;  
 Best Local Similarity 66.7%; Pred. No. 95;

Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GGCGADGPTLRW 12  
 |||||  
 Db 416 GGCGAE-PTREXW 426

## RESULT 25

O8FPC4

ID O8FPC4 PRELIMINARY; PRT; 97 AA.

AC O8FPC4; 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DE Hypothetical protein.  
 GN OrderedlocusNames=CEI185;  
 OS Corynebacterium efficiens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 ON NCBI\_TaxID=152794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN=IS-314;  
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;  
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
 RA Gojobori T.;  
 RT "Comparative complete genome sequence analysis of the amino acid  
 RT replacements responsible for the thermostability of Corynebacterium  
 RT efficiens." 13:1572-1579(2003).  
 RL Genome Res. 13:1572-1579(2003).  
 DR EMBL; AP005220; BAC18668.1; -  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 97 AA; 10632 MW; 6CF1DA566EB304C CRC64;

Query Match 42.2%; Score 46; DB 2; Length 97;  
 Best Local Similarity 58.3%; Pred. No. 26;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCGADGPTLRW 12  
 |||||  
 Db 77 GGALDGRTRKRW 88

RESULT 26

O7MW49 PRELIMINARY; PRT; 117 AA.

AC O7MW49; 01-MAR-2004 (TReMBLrel. 26, Created)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
 DE Hypothetical protein.  
 GN OrderedlocusNames=PG1251;  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Porphyromonadaceae; Porphyromonas.  
 ON NCBI\_TaxID=837;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W83;  
 RX MEDLINE=22829867; PubMed=12949112;  
 DOI=10.1128/JB.185.18.5591-5601.2003;  
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,  
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwim M.L.,  
 RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,  
 RA Granger D., Trevelin H., Dong H., Galvin J.L., Duncan M.J.,  
 RA Dewhirst F.E., Fraser C.M.;  
 RT "Complete genome sequence of the oral pathogenic bacterium  
 RT Porphyromonas gingivalis strain W83.";  
 RL J. Bacteriol. 185:5591-5601(2003).  
 DR EMBL; AF017176; AA066334.1; -  
 TIGR; PG1251; -  
 KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 117 AA; 12589 MW; B4421EB01D186859 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 117;  
 Best Local Similarity 52.9%; Pred. No. 32;  
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GGCGADGPTLRWISFCCG 18  
 |||||  
 Db 17 GCVCVPTVAWITIGAGG 33

## RESULT 27

O8N852 PRELIMINARY; PRT; 159 AA.

AC O8N852; 01-OCT-2002 (TReMBLrel. 22, Created)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
 DE Hypothetical protein FLJ40008.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stomach;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
 RA Sudo H., Hosoliti T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Ninomiya K., Iehibashi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
 RA Kanehori K., Takahashi-Fujii A., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terasahima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukumura Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 DR EMBL; AK097327; BAC04999.1; -  
 SQ SEQUENCE 159 AA; 17782 MW; DF63A4A6D7129A8 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 159;  
 Best Local Similarity 52.6%; Pred. No. 43;  
 Matches 10; Conservative 0; Mismatches 5; Indels 4; Gaps 1;

Qy 1 GGCA-----DQPTLRWISF 15  
 |||||  
 Db 17 GGCGGLVKGKMTLRWASF 35

RESULT 28

O63KH8 PRELIMINARY; PRT; 162 AA.

```

AC 063KX8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BPS31383;
OC Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Tildall R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshaizer D.,
RA Felwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagsels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabbinnowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vearatchavee M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.,
RT "Genomic plasticity of the causative agent of melioidosis,"
RT Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH38655.1; -.
SQ SEQUENCE 162 AA; 17186 MW; 27CDFF4999112A83 CRC64;
QY 2 GCADGPTLR 10
Db 93 GCADGPTLR 101
Query Match 42.2%; Score 46; DB 2; Length 162;
Best Local Similarity 88.9%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GCADGPTLR 10
Db 93 GCADGPTLR 101
RESULT 29
Q7VWMS PRELIMINARY; PRT; 196 AA.
AC Q7VWMS;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocustNames=BP2072;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabbinnowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640417; CAF42350.1; -.
KM Complete proteome; Lipoprotein.
SQ SEQUENCE 196 AA; 21519 MW; FF6E2B8B5EE9968 CRC64;

```

```

Query Match 42.2%; Score 46; DB 2; Length 196;
Best Local Similarity 88.9%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GCADGPTLR 10
Db 18 GCASGPTLR 26
RESULT 30
Q7W9K1 PRELIMINARY; PRT; 196 AA.
AC Q7W9K1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocustNames=BP1756;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabbinnowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640428; CAE37057.1; -.
KM Complete proteome; Lipoprotein.
SQ SEQUENCE 196 AA; 21562 MW; D082FBA6CA6CA765 CRC64;
QY 2 GCADGPTLR 10
Db 18 GCASGPTLR 26
RESULT 31
Q8GVF5 PRELIMINARY; PRT; 245 AA.
AC Q8GVF5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 26, Last annotation update)
DE Putative eukaryotic translation initiation factor 6.
GN Name=OJ1340_C08_131;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC Sasaki T., Matsumoto T., Katayose Y.;
RA "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1340_C08."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```



QY 6 GPTLREWISFCG 18  
 |||||  
 Db 149 GPTLREWLDRVGG 161

## RESULT 35

Q9P858 PRELIMINARY; PRT; 443 AA.  
 AC Q9P858; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein. (Septoria nodorum).  
 OS Phaeosphaeria nodorum (Septoria nodorum).  
 OG Plasmid delta1.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;  
 OC Pleosporales; Phaeosphaeriaceae; Phaeosphaeria.  
 OX NCBI\_TaxID=13684;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BS444;  
 RA "Rawson J.M.;  
 RT "Transposable elements in the phytopathogenic fungus Stagonospora  
 RL nodorum";  
 RN Thesis (2000), PhD thesis, University of Birmingham, UK.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BS444;  
 RA "Rawson J.M., Cutler S.B., Caten C.B.;  
 DL EMBL; AJ277966; CAB91876.1;"  
 RT Hypothetical protein; Plasmid.  
 SQ SEQUENCE 443 AA; 49466 MW; 367E0762EE839E68 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 443;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 CADGPTLREWISFCG 18  
 |||||  
 Db 170 CSENGTLREWITALQG 185

RESULT 36  
 Q6AIT0 PRELIMINARY; PRT; 482 AA.  
 AC Q6AIT0; 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=DP3021;  
 OS Desulfotalea psychrophila.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;  
 OC Desulfobacteriaceae; Desulfobacterae.  
 OX NCBI\_TaxID=84980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LSY54 / DSM 12343;  
 RX PubMed=15305914;  
 RA "Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,  
 RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,  
 RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,  
 RA Klenk H.-P.;  
 RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium  
 RT from permanently cold Arctic sediments.";  
 RL Environ. Microbiol. 6:887-902(2004).  
 DR EMBL; CR522870; CAG37750.1;"  
 DR InterPro; IPR003846; UPF0061.  
 DR Pfam; PF02696; UPF0061; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 482 AA; 54161 MW; 5F401BE2D89323D CRC64;

Query Match 42.2%; Score 46; DB 2; Length 482;  
 Best Local Similarity 69.2%; Pred. No. 1.3e+02;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCADGPTLREWT 13  
 |||||  
 Db 120 GRCVGPALREFI 132

## RESULT 37

AAS5 HUMAN STANDARD; PRT; 926 AA.  
 ID AAS5 HUMAN 09JUR5; 095462;  
 DT 05-JUN-2004 (Rel. 44, Created)  
 DT 05-JUN-2004 (Rel. 44, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alpha-aminoadipic semialdehyde synthase, mitochondrial precursor  
 DE (LKR/SDH) [includes: lysine ketoglutarate reductase (EC 1.5.1.8) (LOR)  
 DE (LKR); Saccharopine dehydrogenase (EC 1.5.1.9) (SDH)].  
 GN Name=AAS5;  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX PubMed=10775527;  
 RA "Sacksteder K.A.; Biery B.J., Morrell J.C., Goodman B.K.,  
 RA Geisbrecht B.V., Cox R.P., Gould S.J., Geraghty M.T.;  
 RT "Identification of the alpha-aminoadipic semialdehyde synthase gene,  
 RT which is defective in familial hyperlysinemia.";  
 RL Am. J. Hum. Genet. 66:1736-1743(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA "Papes F., Kemper E.L., Cord-Neto G., Langone F., Arruda P.;  
 RT "Cloning and expression analysis of the LKR/SDH gene in human  
 RT tissues.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;  
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,  
 RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,  
 RA Wyllie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.B.,  
 RA Fewell G.A., Delehanuy K.D., Miner T.L., Nash W.E., Cordes M., Du H.,  
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,  
 RA Vazirani A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,  
 RA Zerkow P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,  
 RA Strong C.M., Hou S., Tomlinson C., Daughin-Kohlberg S.,  
 RA Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,  
 RA Tin-Mollam A.-M., Abbott A., Minx P., Maupin R., Stromwater C.,  
 RA Lathille P., Miller N., Johnson D., Murray J., Woessner J.P.,  
 RA Wendi M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieghel J.,  
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,  
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,  
 RA Clifton S.W., Chisoe S.L., Marra M.A., Raymond C., Haugen E.,  
 RA Sims E., Zhou Y., James R., Phelps K., Iadonco S., Bubb K.,  
 RA Baertsch R.A., Brent M.R., Kelbler E., Flicek P., Bork P., Suyama M.,  
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,  
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,  
 RA Waterston R.H., Wilson R.K.;  
 RT "The DNA sequence of human chromosome 7.";  
 RL Nature 424:157-164(2003).  
 CC -!- FUNCTION: A bifunctional enzyme that catalyzes the first two steps  
 CC in lysine degradation. The N-terminal and the C-terminal contain  
 CC lysine-ketoglutarate reductase and saccharopine dehydrogenase  
 CC activity, respectively.  
 CC -!- CATALYTIC ACTIVITY: N(6) - (L-1,3-dicarboxypropyl)-L-lysine +  
 CC NADP(+) + H(2)O = L-lysine + 2-oxoglutarate + NADPH.  
 CC -!- CATALYTIC ACTIVITY: N(6) - (L-1,3-dicarboxypropyl)-L-lysine + NAD(+) +  
 CC H(2)O = L-glutamate + 2-aminoadipate 6-semialdehyde + NADH.

```

CC -1- PATHWAY: Lysine degradation; Saccharopine pathway; first step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in all 16 tissues examined with
CC highest expression in the liver.
CC -1- INDUCTION: Induced by starvation (By similarity).
CC -1- DISEASE: Defects in AASS are the cause of hyperlysinemia
CC [MIM:238700]. Hyperlysinemia is an autosomal recessive condition
CC characterized by hyperlysinemia lysinuria and variable
CC saccharopinuria.
CC -1- SIMILARITY: In the N-terminal section; belongs to the ALADH/PNT
CC family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the saccharopine
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF229180; AAF44328.1; -
DR EMBL; AJ007714; CAA07619.2; -
DR EMBL; AC006020; AAF03526.1; -
DR Genew; HGNC:17366; AASS.
DR Reactome; Q9UDR5; -.
DR MIM; 605113; -.
DR MIM; 238700; -.
DR InterPro; IPR007698; Aladh_PNT_C.
DR InterPro; IPR007886; Aladh_PNT_N.
DR InterPro; IPR005097; Saccharopdh.
DR Pfam; PF01262; Aladh_PNT_N; 1.
DR Pfam; PF05222; Aladh_PNT_N; 1.
DR Pfam; PF03435; Saccharopdh; 1.
DR Pfam; PF03435; Multifunctional enzyme; NAD; NADP; Oxidoreductase;
DR Mitochondrion; Multifunctional enzyme; NAD; NADP; Oxidoreductase;
DR TRANSIT 1 32 Mitochondrion (By similarity).
DR CHAIN 33 926 Alpha-aminoacidic semialdehyde synthase.
DR DOMAIN 33 455 Lysine-ketoglutarate reductase.
DR DOMAIN 477 926 Saccharopine dehydrogenase.
DR CONFLICT 589 589 S -> C (in Ref. 2).
DR SEQUENCE 926 AA; 102131 MW; CB4194014351A18D CRC64;
SQ
Query Match 42.2%; Score 46; DB 1; Length 926;
Best Local Similarity 53.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 6 GPTLRWISFPCG 18
DB 623 GATIESYISYCGG 635

```

```

DR EMBL; AF159533; AAD45326.2; -.
DR GO; GO:0000148; C:1,3-beta-glucan synthase complex; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008433; E:1,3-beta-glucan synthase activity; IEA.
DR GO; GO:0006075; P:beta-1,3 glucan biosynthesis; IEA.
DR InterPro; IPR003440; Glyco trans 48.
DR Pfam; PF02364; Glucan synthase; I.
DR SEQUENCE 1902 AA; 217552 MW; 66FC3C60E725F2F CRC64;
SQ
Query Match 42.2%; Score 46; DB 2; Length 1902;
Best Local Similarity 46.7%; Pred. No. 5.1e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 2 GCADGPTLRWISFC 16
DB 1381 GCADINPVADWVQRC 1395

```

RESULT 39

```

ID 08XZNS PRELIMINARY; PRT; 309 AA.
AC 08XZNS;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 26, Last sequence update)
DE PROBABLE TRANSCRIPTION REGULATOR PROTEIN.
OS Name=RS04642; OrderedAccession=RS01360;
GN Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EM11000;
RC MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Gentin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Broctier P., Camus J.C., Cartolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavié M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646064; CAD15062.1; -.
DR HSSP; GQWXC7; IIXC.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00126; HTH 1; 1.
DR Pfam; PF03466; LysR substructure; 1.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 33774 MW; 733551741CE83182 CRC64;

```

Query Match 41.7%; Score 45.5; DB 2; Length 309;

Best Local Similarity 60.0%; Pred. No. 99;

Matches 9; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1 GG---CADGPTLRW 12

DB 216 GGTMECTDGAVALREW 230

RESULT 40

```

ID 07RUA5 PRELIMINARY; PRT; 108 AA.
AC 07RUA5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein B24B19.30.
GN Name=NCU03933.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

```

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=0874a;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30070D15 product:unclassified, full insert sequence.  
 GN Name=C130070B15R1k;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Meth. Enzymol. 303:19-44(1999).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsunoto H., Sakaiguchi S., Ikegami T., Kaishiwagi K., Fujiwaka S., Inoue K., Togawa Y., Ikawa M., Ohara E., Watanabe M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Akimura T., Hara A., Hashizume W., Furuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ono M., Ohata N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

RESULT 41  
 ID 06ZTT4 PRELIMINARY; PRT; 146 AA.  
 AC 06ZTT4;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ44235.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S., Tanai H., Watanabe S., Ishida S., Ono Y., Houta T., Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuo Y., Nagai K., Isogai T.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK126223; BAC6495.1 -  
 SQ SEQUENCE 146 AA; 16475 MW; COB7BBE49151B89B CRC64;  
 Query Match 41.3%; Score 45; DB 2; Length 108;  
 Best Local Similarity 50.0%; Pred. No. 42;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Qy 3 CADGPTLEWISFC 16  
 Db 70 CQCQPIRLNMLSWC 83

Qy 2 GCADGPTLEWIS 14  
 Db 28 GCADGCVLRGRYS 40  
 RESULT 42



RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK081706; BAC38302.1; -;  
 DR MGD; MG:2444974; C130070B1SRK.  
 SQ SEQUENCE 173 AA; 19340 MW; 6227DD6725852PCD CRC64;

Query Match 41.3%; Score 45; DB 2; Length 173;  
 Best Local Similarity 63.6%; Pred. No. 67;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 GPTLRWISFC 16  
 Db 75 GVTREMASWC 85

RESULT 43  
 007291 PRELIMINARY; PRT; 180 AA.

AC 007291;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein orfs.  
 GN Name-orfs;  
 OS Natronomonas pharaonis (Natronobacterium pharaonis).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Natronomonas.  
 OX NCBI\_TaxID=2257;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SP1/28;  
 RC MEDLINE=98088958; PubMed=9428682;  
 RA Matlar S., Engelhard M.;  
 RT "Cytochrome b3 from Natronobacterium pharaonis: An archaeal four-subunit cytochrome-c-type oxidase.";  
 RL Eur. J. Biochem. 250:332-341(1997).  
 DR EMBL; Y10500; CAA71527.1; -;  
 DR PIR; T44944; T44944.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR010985; Met\_repress\_like.  
 DR K0 Hypothetical protein  
 SQ SEQUENCE 180 AA; 20215 MW; A8C3739BC8C11310 CRC64;

Query Match 41.3%; Score 45; DB 2; Length 180;  
 Best Local Similarity 77.8%; Pred. No. 69;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 LREWISFCG 17  
 Db 116 LLEWLSFCG 124

RESULT 44  
 06N1X5 PRELIMINARY; PRT; 209 AA.

AC 06N1X5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedlocusNames=RP44277;  
 OS Rhodopseudomonas palustris.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Rhodopseudomonas.  
 OX NCBI\_TaxID=1076;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGA009 / ATCC BAA-98;  
 RC PubMed=14704707; DOI=10.1038/nbc923;  
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,  
 RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,  
 RA Gibson J.L., Hanson T.E., Bobst C., Torres Y., Torres J.L., Perez C.,  
 RA Harrison F.H., Gibson J., Harwood C.S.;

RT "Complete genome sequence of the metabolically versatile  
 RT photosynthetic bacterium Rhodopseudomonas palustris.";  
 RL Nat. Biotechnol. 22:55-61(2004).  
 DR EMBL; BX572606; CAE29718.1; -;  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF02985; HEAT; 2.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84DB040EE CRC64;

Query Match 41.3%; Score 45; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 81;  
 Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 3 CADG-----PTLRWIS 14  
 Db 98 CADTGYEALPTLRWLS 115

RESULT 45

089JRS PRELIMINARY; PRT; 290 AA.

AC 089JRS;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE B115207 protein.  
 GN OrderedlocusNames=b115207;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USD110;  
 RC MEDLINE=22484998; PubMed=12597275;  
 RX Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
 RA Sasamoto S., Katanahe A., Iidesawa K., Iriguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpō S., Tsurunaka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AF005854; BAC50472.1; -;  
 DR EMBL; AF005854; BAC50472.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 290 AA; 30111 MW; CAE33930D6CFC7FF CRC64;

Query Match 41.3%; Score 45; DB 2; Length 290;  
 Best Local Similarity 61.5%; Pred. No. 11e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 DGPTLRWISFCG 17  
 Db 269 DGAPLEWIAFAg 281

Search completed: September 1, 2005, 16:20:56  
 Job time : 72.9496 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 87.3453 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-7

Perfect score: 108  
Sequence: 1 GNADGPTLRQWLEGRPKN 19

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	108	100.0	19 2 AAW09457	AAW09457 Thrombopo
2	108	100.0	19 2 AAW09492	AAW09492 Thrombopo
3	108	100.0	19 2 AAW36651	AAW36651 Thrombopo
4	108	100.0	19 2 AAW33024	AAW33024 Thrombopo
5	108	100.0	19 2 AAW36643	AAW36643 Thrombopo
6	108	100.0	19 3 AAB17021	AAB17021 TPO-mimet
7	108	100.0	19 4 AAW25862	AAW25862 Human thr
8	108	100.0	19 4 AAW25870	AAW25870 Human thr
9	108	100.0	19 4 AAW25821	AAW25821 Human thr
10	108	100.0	19 5 ABB72907	ABB72907 TPO mimet
11	108	100.0	19 7 ADJ73059	ADJ73059 TPO mimet
12	108	100.0	19 8 ADJ52694	ADJ52694 CHI delet
13	108	100.0	19 8 ADJ51655	ADJ51655 CHI delet
14	108	100.0	19 8 ADJ51687	ADJ51687 TPO mimet
15	62	57.4	18 5 ABB73389	ABB73389 TPO-mimet
16	62	57.4	18 8 ADQ16617	ADQ16617 TPO mimet
17	62	57.4	22 7 ADNS59822	ADNS59822 TMP pepti
18	62	57.4	22 7 ADNS59792	ADNS59792 Peptide-v
19	62	57.4	23 7 ADNS59774	ADNS59774 Peptide-v
20	62	57.4	25 7 ADNS59692	ADNS59692 Thrombopo
21	62	57.4	36 7 ADNS59762	ADNS59762 Peptide-v
22	62	57.4	41 7 ADNS59768	ADNS59768 Peptide-v
23	62	57.4	43 7 ADNS59761	ADNS59761 Peptide-v
24	62	57.4	44 7 ADNS59817	ADNS59817 Peptide-v
25	62	57.4	46 7 ADNS59780	ADNS59780 Peptide-v

26	62	57.4	46 7 ADNS59786	ADNS59786 Peptide-v
27	60	55.6	18 7 ADNS59667	ADNS59667 Thrombopo
28	60	55.6	22 7 ADNS59834	ADNS59834 TMP pepti
29	60	55.6	25 7 ADNS59716	ADNS59716 Thrombopo
30	60	55.6	34 3 AAY96527	AAY96527 Thrombopo
31	60	55.6	42 7 ADNS59818	ADNS59818 Peptide-
32	59	54.6	15 5 ADNS51670	ADNS51670 Thrombopo
33	59	54.6	15 5 ADQ16585	ADQ16585 TPO mimet
34	59	54.6	18 5 ADNS51689	ADNS51689 TPO mimet
35	59	54.6	18 5 ADNS51688	ADNS51688 TPO mimet
36	59	54.6	18 5 ADNS51686	ADNS51686 TPO mimet
37	59	54.6	18 5 ADNS51693	ADNS51693 TPO mimet
38	59	54.6	18 5 ADNS51684	ADNS51684 TPO mimet
39	59	54.6	18 5 ADNS51691	ADNS51691 TPO mimet
40	59	54.6	18 5 ADNS51690	ADNS51690 TPO mimet
41	59	54.6	18 5 ADNS51675	ADNS51675 TPO mimet
42	59	54.6	18 5 ADQ16611	ADQ16611 TPO mimet
43	59	54.6	18 8 ADQ16619	ADQ16619 TPO mimet
44	59	54.6	18 8 ADQ16621	ADQ16621 TPO mimet
45	59	54.6	18 8 ADQ16646	ADQ16646 TPO mimet
46	59	54.6	18 8 ADQ16615	ADQ16615 TPO mimet
47	59	54.6	18 8 ADQ16625	ADQ16625 TPO mimet
48	59	54.6	18 8 ADQ16629	ADQ16629 TPO mimet
49	59	54.6	18 8 ADQ16623	ADQ16623 TPO mimet
50	59	54.6	22 8 ADQ16710	ADQ16710 Immunog10
51	59	54.6	128 8 ADQ16705	ADQ16705 Modified
52	59	54.6	225 8 ADQ16704	ADQ16704 Modified
53	59	54.6	472 5 ADNS51695	ADNS51695 5G1.1-TPO
54	59	54.6	472 8 ADQ16647	ADQ16647 Immunog10
55	58	53.7	18 3 AAB16957	AAB16957 PBEGylated
56	58	53.7	18 3 AAB16956	AAB16956 PBEGylated
57	58	53.7	18 5 ADNS51674	ADNS51674 TPO mimet
58	58	53.7	18 5 ADNS51683	ADNS51683 TPO mimet
59	58	53.7	18 8 ADQ16607	ADQ16607 TPO mimet
60	58	53.7	18 8 ADQ16609	ADQ16609 TPO mimet
61	58	53.7	19 5 ABB73390	ABB73390 TPO-mimet
62	58	53.7	20 3 AAB18003	AAB18003 Fc-TMP pe
63	58	53.7	20 5 ABB73403	ABB73403 TPO mimet
64	58	53.7	22 8 ADQ16709	ADQ16709 Immunog10
65	58	53.7	30 3 AAB17287	AAB17287 TPO-mimet
66	58	53.7	31 3 AAB17288	AAB17288 TPO-mimet
67	58	53.7	32 3 AAB17289	AAB17289 TPO-mimet
68	58	53.7	33 3 AAB17290	AAB17290 TPO-mimet
69	58	53.7	34 3 AAB17291	AAB17291 TPO-mimet
70	58	53.7	35 3 AAB17292	AAB17292 TPO-mimet
71	58	53.7	36 3 AAY96525	AAY96525 Thrombopo
72	58	53.7	36 3 AAY96523	AAY96523 Thrombopo
73	58	53.7	36 3 AAY96524	AAY96524 Thrombopo
74	58	53.7	36 3 AAY96526	AAY96526 Thrombopo
75	58	53.7	36 3 AAB17307	AAB17307 TPO-mimet
76	58	53.7	36 3 AAB17293	AAB17293 TPO-mimet
77	58	53.7	36 3 AAB17303	AAB17303 TPO-mimet
78	58	53.7	36 3 AAB16963	AAB16963 TPO-mimet
79	58	53.7	36 3 AAB17301	AAB17301 TPO-mimet
80	58	53.7	36 3 AAB17306	AAB17306 TPO-mimet
81	58	53.7	36 5 ABB72403	ABB72403 TPO-mimet
82	58	53.7	37 3 AAB17294	AAB17294 TPO-mimet
83	58	53.7	38 3 AAB17295	AAB17295 TPO-mimet
84	58	53.7	39 3 AAB17304	AAB17304 TPO-mimet
85	58	53.7	39 3 AAB17305	AAB17305 TPO-mimet
86	58	53.7	40 3 AAB17302	AAB17302 TPO-mimet
87	58	53.7	41 3 AAY96528	AAY96528 Thrombopo
88	58	53.7	41 5 ABB73389	ABB73389 TPO-mimet
89	58	53.7	41 5 ABB73388	ABB73388 TPO-mimet
90	58	53.7	42 3 AAY96530	AAY96530 TPO-mimet
91	58	53.7	42 3 AAB17296	AAB17296 Thrombopo
92	58	53.7	42 3 AAB17308	AAB17308 Synthetic
93	58	53.7	42 3 AAB17282	AAB17282 TPO-mimet
94	58	53.7	42 3 AAB17281	AAB17281 TPO-mimet
95	58	53.7	42 5 ABB73404	ABB73404 TMP-TMP g
96	58	53.7	60 3 AAB17311	AAB17311 Synthetic
97	58	53.7	60 5 ABB73405	ABB73405 TMP-TMP g
98	58	53.7	247 3 AAB16958	AAB16958 Fc-TMP pr



Db 1 GNAAGPTLRQWLBSGRPRKN 19

## RESULT 3

AAW36651 ID AAW36651 standard; peptide; 19 AA.

XX AAW36651;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 27; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX Sequence 19 AA;

Query Match 100.0%; Score 108; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.1e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAAGPTLRQWLBSGRPRKN 19

Db 1 GNAAGPTLRQWLBSGRPRKN 19

## RESULT 4

AAW33024 ID AAW33024 standard; peptide; 19 AA.

XX AAW33024;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX

KW Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a

CC molecular weight of less than 8000 Da and a TR binding affinity as

CC expressed by an IC50 of no more than about 100 microm. It can be used to

CC treat disorders which are susceptible to treatment with a thrombopoietin

CC agonist, preferably haematological disorders and thrombocytopenia

CC resulting from chemotherapy, radiation therapy or bone marrow

CC transfusions. It can also be used diagnostically, e.g. to investigate the

CC mechanism of thrombopoietin signal transduction and receptor activation,

CC or to maintain the proliferation and growth of thrombopoietin dependent

XX cell lines

XX Sequence 19 AA;

Query Match 100.0%; Score 108; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.1e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNAAGPTLRQWLBSGRPRKN 19

Db 1 GNAAGPTLRQWLBSGRPRKN 19

## RESULT 5

AAW36643 ID AAW36643 standard; peptide; 19 AA.

XX AAW36643;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS,  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 PI WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Disclosure; Page 26; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 108; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRRPKN 19  
 1 GNADGPTLRQWLEGRRPKN 19  
 Db 1 GNADGPTLRQWLEGRRPKN 19  
 RESULT 6  
 AAB17021  
 ID AAB17021 standard; peptide; 19 AA.  
 AC  
 XX AAB17021;  
 XX  
 DT 31-OCT-2000 (first entry)  
 XX  
 DE TPO-mimetic peptide sequence SEQ ID NO:77.  
 XX  
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; Epo; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200024782-A2.  
 XX  
 PD 04-MAY-2000.  
 XX  
 PF 25-OCT-1999; 99WO-US025044.  
 XX  
 PR 23-OCT-1998; 98US-0105371P.  
 PR 22-OCT-1999; 99US-00428082.  
 XX  
 PA (AMGEN-) AMGEN INC.  
 XX  
 PI Feige U, Liu C, Cheetham J, Boone TC;  
 PI WPI; 2000-350702/30.  
 DR  
 XX Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 XX Claim 19; Page 220; 608pp; English.  
 PS  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-  
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement functions  
 CC possibly placental transfer. AA69443 to AA69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 108; DB 3; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRRPKN 19  
 1 GNADGPTLRQWLEGRRPKN 19  
 Db 1 GNADGPTLRQWLEGRRPKN 19  
 RESULT 7  
 AAU25862  
 ID AAU25862 standard; peptide; 19 AA.  
 AC  
 XX AAU25862;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #48.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US008623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Depirince RB, Poddaturi S;  
 PI Yin Q;  
 PI WPI; 2001-564142/63.  
 DR  
 XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX Sequence 19 AA;  
 SQ  
 Query Match 100.0%; Score 108; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRRPKN 19  
 Db 1 GNADGPTLRQWLEGRRPKN 19  
 RESULT 8  
 AAU25870  
 ID AAU25870 standard; peptide; 19 AA.  
 XX  
 AC AAU25870;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #56.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S,  
 PI Yin Q;  
 XX WPI; 2001-564142/63.  
 XX

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX Sequence 19 AA;  
 SQ  
 Query Match 100.0%; Score 108; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRRPKN 19  
 Db 1 GNADGPTLRQWLEGRRPKN 19  
 RESULT 9  
 AAU25821  
 ID AAU25821 standard; peptide; 19 AA.  
 XX  
 AC AAU25821;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #7.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S,  
 PI Yin Q;  
 XX WPI; 2001-564142/63.  
 XX

XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 65-66; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptide and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX  
 SQ Sequence 19 AA;  
 XX  
 Query Match 100.0%; Score 108; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRPRKN 19  
 ID 1 GNADGPTLRQWLEGRPRKN 19  
 DB ABB72907 standard; peptide; 19 AA.  
 XX  
 AC ABB72907;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:77.  
 XX  
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNF;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antineumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antinaemic; anorectic; antifertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic; retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200183525-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 02-MAY-2001; 2001WO-US014310.  
 XX  
 PR 03-MAY-2000; 2000US-00563286.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudae JW;  
 XX  
 DR WPI; 2002-130313/17.  
 XX  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 44; 176pp; English.  
 XX  
 CC The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antineumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antinaemic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB72426 and ABL35695 to ABB35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 19 AA;  
 XX  
 Query Match 100.0%; Score 108; DB 5; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRPRKN 19  
 ID 1 GNADGPTLRQWLEGRPRKN 19  
 DB 1 GNADGPTLRQWLEGRPRKN 19  
 XX  
 AC 1 GNADGPTLRQWLEGRPRKN 19  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE TPO mimetic peptide sequence SEQID 513.  
 XX  
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
 KW TPO.  
 XX  
 OS Synthetic.  
 OS  
 PN WO2003084477-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 24-MAR-2003; 2003WO-US009139.  
 XX  
 PR 29-MAR-2002; 2002US-0368791P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;  
 XX

DR WPI; 2003-804237/75.  
XX  
PT New CDR mimetibody comprising a portion of a heavy or light chain  
PT variable region comprising human framework or ligand binding region,  
PT useful for preparing a composition for treating e.g., immune,  
PT cardiovascular or neurologic disease.  
XX  
PS Disclosure; SEQ ID NO 513; 97pp; English.  
XX  
CC This invention relates to novel mammalian CDR mimetibodies, specific  
CC portions or variants thereof. Specifically, it refers to an antibody  
CC fragment where a protein has been inserted into, or replaces a portion  
CC of one or more CDR regions, such that each CDR mimetibody comprises at  
CC least one portion of a heavy chain or light chain variable region, which  
CC itself comprises at least one human framework region and at least one  
CC ligand binding region (LBR). The present invention describes human  
CC mimetibodies, including modified immunoglobulins and cleavage products  
CC that can be useful in gene therapy and the generation of transgenic  
CC plants and animals. Furthermore, the CDR mimetibody is useful for  
CC preparing compositions for modulating, treating or reducing the symptoms  
CC of immune, cardiovascular, infectious, malignant and/or neurologic  
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
CC peptide sequence is a TPO mimetic peptide sequence used to make a  
CC mimetibody of the invention.  
XX  
SQ Sequence 19 AA;  
XX  
Query Match 100.0%; Score 108; DB 7; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 GNADGPTLRQWLEGRPRKN 19  
DB 1 GNADGPTLRQWLEGRPRKN 19  
XX  
RESULT 12  
ADJ52694  
ID ADJ52694 standard; peptide; 19 AA.  
XX  
AC ADJ52694;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE CH1 deleted mimetibody-related peptide SeqID513.  
XX  
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
KW hypotensive; neuroprotective; nootropic; antibacterial; virocidic;  
KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
KW arrhythmia; hypertension; heart failure; neurodegenerative;  
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
KW cancerous condition; infectious disease; bacterial infection;  
KW viral infection; fungal infection.  
XX  
OS Unidentified.  
OS Synthetic.  
XX  
FN WO2004002417-A2.  
XX  
PD 08-JAN-2004.  
XX  
XX 27-JUN-2003; 2003WO-US020347.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX  
XX Heavner GA, Knight DM, Ghreyeb J, Scallion BJ, Neseppor TC,  
XX Kucolowski KA;  
XX WPI; 2004-082870/08.  
XX

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
PT diseases.  
XX  
PS Claim 2; SEQ ID NO 513; 129pp; English.  
XX  
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an immunosuppressive,  
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
CC antibacterial, virocidic or fungicide activity. In addition, the disclosed  
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
CC is useful for diagnosing or treating a disease condition in a cell,  
CC tissue, organ or animal, specifically for modulating, treating,  
CC alleviating, preventing the incidence or reducing the symptoms of an  
CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
CC conditions, or infectious diseases (for example bacterial, viral or  
CC fungal infection). The present sequence is that of a peptide which may be  
CC used during the creation of a mimetibody of the invention.  
XX  
SQ Sequence 19 AA;  
XX  
Query Match 100.0%; Score 108; DB 8; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 GNADGPTLRQWLEGRPRKN 19  
DB 1 GNADGPTLRQWLEGRPRKN 19  
XX  
RESULT 13  
ADJ51655  
ID ADJ51655 standard; peptide; 19 AA.  
XX  
AC ADJ51655;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE CH1 deleted mimetibody-related peptide SeqID513.  
XX  
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
KW obstructive disorder; haematologic disorder; immunologic disorder;  
KW bacterial disorder; infectious disorder; musculoskeletal disorder;  
KW allergic disorder; neurological disorder; nutritional disorder;  
KW oncological disorder; neurological disorder; nutritional disorder;  
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
KW renal disorder; pulmonary disorder.  
XX  
OS Unidentified.  
OS Synthetic.  
XX  
FN WO2004002424-A2.  
XX  
XX 08-JAN-2004.  
XX  
XX 30-JUN-2003; 2003WO-US020495.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX 19-SEP-2002; 2002US-0412144P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX



PI Heavener GA, Knight DM, Ghayab J, Scallion BJ, Nesspor TC;  
 PI Kutolowski KA;  
 XX  
 DR WPI, 2004-082872/08.  
 XX  
 PT New CHI deleted mimetic peptide polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 15, SEQ ID NO 513; 123pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetic bodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-gen, dermatological-gen, auditory, endocrine-gen,  
 CC gastrointestinal-gen, gynaecological-gen, hepatotropic, haemostatic,  
 CC immunomodulator, anti-allergic, muscular-gen, cytostatic,  
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetic body, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstructive, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetic body of the invention.  
 XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 108; DB 8; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRPPK 19  
 Db 1 GNADGPTLRQWLEGRPPK 19

XX  
 DR WPI, 2002-566610/60.  
 DR N-PSDB; ABQ73365.  
 XX  
 PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.  
 XX  
 PS Claim 20; Fig 5; 113pp; English.  
 XX  
 CC The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (Epo) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;  
 Query Match 57.4%; Score 62; DB 5; Length 18;  
 Best Local Similarity 64.7%; Pred. No. 0.016;  
 Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRPP 17  
 Db 1 GRIGGPTLRQWLAAAP 17

RESULT 14  
 ABP51687  
 ID ABP51687 standard; peptide: 18 AA.  
 XX  
 AC ABP51687;  
 XX  
 DT 01-OCT-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:37.  
 XX  
 KW TPO, EPO, thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200246238-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 05-DEC-2001; 2001WO-US047656.  
 XX  
 PR 05-DEC-2000; 2000US-0251448P.  
 PR 04-MAY-2001; 2001US-028889P.  
 PR 29-MAY-2001; 2001US-0294068P.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

RESULT 15  
 ADN59655  
 ID ADN59655 standard; peptide: 18 AA.  
 XX  
 AC ADN59655;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Thrombopoietin mimetic peptide (TMP4), seq id 4.  
 XX  
 KW Haemostatic; antianaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KW thrombocytopoenaia; aplastic anaemia; autoimmune thrombocytopoenaia;  
 KW autoimmune haemolytic anaemia; Hughes' s syndrome;  
 KW lupoid thrombocytopoenaia.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2003031589-A2.  
 XX  
 PD 17-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032552.  
 XX  
 PR 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX



PA (AMGE-) AMGEN INC.  
 XX  
 PI Min H, Sitney KC, Hartley C;  
 XX  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 XX  
 PS Claim 6; SEQ ID NO 4; 126pp; English.  
 XX  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (i) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (ii) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (ii) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a preferred TMP of the invention.  
 CC  
 SQ Sequence 18 AA;  
 XX  
 Query Match 57.4%; Score 62; DB 7; Length 18;  
 Best Local Similarity 84.6%; Pred. No. 0.016;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLEGRR 16  
 |||||:|||||  
 Db 4 DGPTLRQWLEYYR 16  
 |||||:|||||  
 RESULT 16  
 ADO16617  
 ID ADO16617 standard; peptide; 18 AA.  
 XX  
 AC ADO16617;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE TPO mimetic peptide with random flanking residues SEQ ID NO:37.  
 XX  
 KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.  
 XX  
 OS Unidentified.  
 OS  
 PN WO2004050017-A2.  
 PD 17-JUN-2004.  
 XX  
 PF 17-NOV-2003; 2003WO-US036894.  
 XX  
 PR 02-DEC-2002; 2002US-00307724.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.

XX  
 PI Bowdish KS, Frederickson S, Renshaw M;  
 XX  
 DR WPI; 2004-460973/43.  
 DR N-PSDB; ADO16618.  
 XX  
 PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX  
 PS Example 1; SEQ ID NO 37; 107pp; English.  
 XX  
 CC The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.  
 CC  
 SQ Sequence 18 AA;  
 XX  
 Query Match 57.4%; Score 62; DB 8; Length 18;  
 Best Local Similarity 64.7%; Pred. No. 0.016;  
 Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 GNADGPTLRQWLEGRR 17  
 :|||||:  
 Db 1 GPISGPTLRQWLARAP 17  
 |||||:  
 RESULT 17  
 ADN59822  
 ID ADN59822 standard; peptide; 22 AA.  
 XX  
 AC ADN59822;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE TMP peptide TMP4.  
 XX  
 KW Haemostatic; antihaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KW autoimmune haemolytic anaemia; Hughes's syndrome;  
 KW lupoid thrombocytopenia; linker.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2003031589-A2.  
 PD 17-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032552.  
 XX  
 PR 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Min H, Sitney KC, Hartley C;  
 XX  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 XX

PS Example 6; Page 83; 126pp; English.  
 XX  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocyte or  
 CC platelet in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a TMP peptide of the invention to which a two amino acid "cap"  
 CC has been added to the carboxy terminal to increase peptide affinity.  
 XX  
 SQ Sequence 22 AA;

Query Match 57.4%; Score 62; DB 7; Length 22;  
 Best Local Similarity 84.6%; Pred. No. 0.02;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 DGEPTLRQMLEGRR 16  
 |||||:||||  
 Db 6 DGEPTLRQMLEYRR 18

## RESULT 18

ADN59792  
 ID ADN59792 standard; protein; 23 AA.

AC ADN59792;

DT 01-JUL-2004 (first entry)

DE Peptide-vehicle compound, seq id 144.

XX Haemostatic; antihaemic; immunosuppressive; platelet;  
 KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KM thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KM autoimmune haemolytic anaemia; Hughes' syndrome;  
 KM lupoid thrombocytopenia.

XX Unidentified.

OS  
 PN WO2003031589-A2.

PD 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

PA Min H, Sitney KC, Hartley C;

PI WPI; 2003-403101/38.

XX

PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.

PS Disclosure; SEQ ID NO 144; 126pp; English.

XX  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a peptide-vehicle compound of the invention.  
 XX  
 SQ Sequence 23 AA;

Query Match 57.4%; Score 62; DB 7; Length 23;  
 Best Local Similarity 84.6%; Pred. No. 0.021;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 DGEPTLRQMLEGRR 16  
 |||||:||||  
 Db 4 DGEPTLRQMLEYRR 16

## RESULT 19

ADN59774  
 ID ADN59774 standard; protein; 23 AA.

AC ADN59774;

DT 01-JUL-2004 (first entry)

DE Peptide-vehicle compound, seq id 126.

XX Haemostatic; antihaemic; immunosuppressive; platelet;  
 KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KM thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KM autoimmune haemolytic anaemia; Hughes' syndrome;  
 KM lupoid thrombocytopenia.

XX Unidentified.

OS  
 PN WO2003031589-A2.

PD 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

PA Min H, Sitney KC, Hartley C;

PI

XX WPI; 2003-403101/38.

DR Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

XX PT which stimulate the production of platelets and/or the production of

PT platelet precursors, useful for treating thrombocytopenia.

XX

PS Disclosure; SEQ ID NO 126; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that

CC binds to the c-mpl (mpl) receptor, and which stimulates the production of

CC platelets and/or the production of platelet precursors, is new. Further

CC disclosed is a composition of matter (II) that binds to an mpl receptor,

CC and a pharmaceutical composition comprising (II) and a carrier. The

CC pharmaceutical composition of the invention is useful for treating

CC thrombocytopenia in an animal, and for increasing megakaryocytes or

CC platelets in a patient. The TMP of the invention is useful for treating

CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.

CC disease conditions involving thrombocytopenia such as aplastic anaemia,

CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,

CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid

CC thrombocytopenia. The TMP of the invention is also useful for

CC maintaining the viability or storage life of platelets and/or

CC megakaryocytes and its derived cells. The compounds demonstrate an

CC improved ability to bind to and/or trigger transmembrane signal through,

CC i.e. activating, the mpl receptor the compounds have superior

CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in

CC vitro, the production of platelets and/or megakaryocytic activity,

CC i.e. the ability to stimulate, in vivo and in vitro, the production of

CC platelet precursors. Further, certain of the compounds also exhibit

CC superior therapeutic properties, such as improved plasma half-life,

CC biological activity and in vivo circulation time. The current sequence

CC represents a peptide-vehicle compound of the invention.

XX

SQ Sequence 23 AA;

Query Match 57.4%; Score 62; DB 7; Length 23;

Best Local Similarity 84.6%; Pred. No. 0.021;

Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 DGPTLRQWLEGR 16

Db 9 DGPTLRQWLEGR 21

RESULT 20

ADNS9692

ID ADNS9692 standard; peptide; 25 AA.

XX

AC ADNS9692;

XX

DT 01-JUL-2004 (first entry)

XX

DE Thrombopoietin mimetic peptide TMP4, seq id 41.

XX

KM Haemostatic; antihaemic; immunosuppressive; platelet;

KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;

KM thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KM autoimmune haemolytic anaemia; Hughes' syndrome;

KM lupoid thrombocytopenia.

XX

OS Homo sapiens.

XX

PN WO2003031589-A2.

XX

PD 17-APR-2003.

XX

PF 11-OCT-2002; 2002WO-US032552.

XX

PR 11-OCT-2001; 2001US-0328666P.

XX

PR 10-OCT-2002; 2002US-00269806.

XX

PA (AMGE-) AMGEN INC.

XX

PI Min H, Sitney KC, Hartley C;

XX

DR WPI; 2003-403101/38.

XX

DR N-PSDB; ADNS9690.

XX

PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

PT which stimulate the production of platelets and/or the production of

PT platelet precursors, useful for treating thrombocytopenia.

XX

PS Disclosure; SEQ ID NO 41; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that

CC binds to the c-mpl (mpl) receptor, and which stimulates the production of

CC platelets and/or the production of platelet precursors, is new. Further

CC disclosed is a composition of matter (II) that binds to an mpl receptor,

CC and a pharmaceutical composition comprising (II) and a carrier. The

CC pharmaceutical composition of the invention is useful for treating

CC thrombocytopenia in an animal, and for increasing megakaryocytes or

CC platelets in a patient. The TMP of the invention is useful for treating

CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.

CC disease conditions involving thrombocytopenia such as aplastic anaemia,

CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,

CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid

CC thrombocytopenia. The TMP of the invention is also useful for

CC maintaining the viability or storage life of platelets and/or

CC megakaryocytes and its derived cells. The compounds demonstrate an

CC improved ability to bind to and/or trigger transmembrane signal through,

CC i.e. activating, the mpl receptor the compounds have superior

CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in

CC vitro, the production of platelets and/or megakaryocytic activity,

CC i.e. the ability to stimulate, in vivo and in vitro, the production of

CC platelet precursors. Further, certain of the compounds also exhibit

CC superior therapeutic properties, such as improved plasma half-life,

CC biological activity and in vivo circulation time. The current sequence

CC represents a TMP fragment.

XX

SQ Sequence 25 AA;

Query Match 57.4%; Score 62; DB 7; Length 25;

Best Local Similarity 84.6%; Pred. No. 0.022;

Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 DGPTLRQWLEGR 16

Db 7 DGPTLRQWLEGR 19

RESULT 21

ADNS9762

ID ADNS9762 standard; protein; 36 AA.

XX

AC ADNS9762;

XX

DT 01-JUL-2004 (first entry)

XX

DE Peptide-vehicle compound, seq id 114.

XX

KM Haemostatic; antihaemic; immunosuppressive; platelet;

KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;

KM thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KM autoimmune haemolytic anaemia; Hughes' syndrome;

KM lupoid thrombocytopenia.

XX

OS unidentified.

XX

PN WO2003031589-A2.

XX

PD 17-APR-2003.

XX

PF 11-OCT-2002; 2002WO-US032552.

XX

XX 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX  
 PA (AMGE-) AMGEN INC.  
 PI Min H, Sitney KC, Hartley C;  
 XX WPI; 2003-403101/38.  
 DR  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopaenia.  
 PS  
 XX Disclosure; SEQ ID NO 114; 126pp; English.  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmacological composition of the invention is useful for treating  
 CC thrombocytopaenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
 CC disease conditions involving thrombocytopaenia such as aplastic anaemia,  
 CC autoimmune haemolytic anaemia, drug induced immune thrombocytopaenia,  
 CC thrombocytopaenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a peptide-vehicle compound of the invention.  
 CC  
 SQ Sequence 36 AA;  
 Query Match 57.4%; Score 62; DB 7; Length 36;  
 Best Local Similarity 84.6%; Pred. No. 0.033;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 DGPTLKQWLEGR 16  
 Db 4 DGPTLKQWLEGR 16  
 RESULT 22  
 ID ADN59768 standard; protein; 41 AA.  
 AC ADN59768;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Peptide-vehicle compound, seq id 120.  
 XX  
 XX Haemostatic; antihaemic; immunosuppressive; platelet;  
 KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KM thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;  
 KM autoimmune haemolytic anaemia; Hughe's syndrome;  
 KM lupoid thrombocytopaenia.  
 XX  
 OS Unidentified.  
 PN W02003031589-A2.  
 XX

PD 17-APR-2003.  
 XX  
 XX 11-OCT-2002; 2002WO-US032552.  
 PF  
 XX  
 PR 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX  
 PA (AMGE-) AMGEN INC.  
 PI Min H, Sitney KC, Hartley C;  
 XX WPI; 2003-403101/38.  
 DR  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopaenia.  
 PS  
 XX Disclosure; SEQ ID NO 120; 126pp; English.  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmacological composition of the invention is useful for treating  
 CC thrombocytopaenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
 CC disease conditions involving thrombocytopaenia such as aplastic anaemia,  
 CC autoimmune haemolytic anaemia, drug induced immune thrombocytopaenia,  
 CC thrombocytopaenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a peptide-vehicle compound of the invention.  
 CC  
 SQ Sequence 41 AA;  
 Query Match 57.4%; Score 62; DB 7; Length 41;  
 Best Local Similarity 84.6%; Pred. No. 0.038;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 DGPTLKQWLEGR 16  
 Db 4 DGPTLKQWLEGR 16  
 RESULT 23  
 ID ADN59761 standard; protein; 43 AA.  
 AC ADN59761;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Peptide-vehicle compound, seq id 113.  
 XX  
 XX Haemostatic; antihaemic; immunosuppressive; platelet;  
 KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KM thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;  
 KM autoimmune haemolytic anaemia; Hughe's syndrome;  
 KM lupoid thrombocytopaenia.  
 XX  
 OS Unidentified.  
 XX

XX WO2003031589-A2.  
 PN 17-APR-2003.  
 PD 11-OCT-2002; 2002WO-US032552.  
 XX 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 PA (AMGEN-) AMGEN INC.  
 PI Min H, Sitney KC, Hartley C;  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 PS Disclosure; SEQ ID NO 113; 126pp; English.  
 XX  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a peptide-vehicle compound of the invention.  
 XX  
 SQ Sequence 43 AA;  
 Query Match 57.4%; Score 62; DB 7; Length 43;  
 Best Local Similarity 84.6%; Pred. No. 0.041;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLKGR 16  
 DB 4 DGPTLRQWLKGR 16  
 ID ADN59817 standard; peptide; 44 AA.  
 AC ADN59817;  
 DT 01-JUL-2004 (first entry)  
 XX  
 DB Peptide-linker compound, seq id 101.  
 XX  
 KW Haemostatic; antihaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KW autoimmune haemolytic anaemia; Hughes's syndrome;

KW lupoid thrombocytopenia; linker.  
 XX  
 OS unidentified.  
 XX  
 PN WO2003031589-A2.  
 XX 17-APR-2003.  
 PD 11-OCT-2002; 2002WO-US032552.  
 XX 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 PA (AMGEN-) AMGEN INC.  
 PI Min H, Sitney KC, Hartley C;  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 PS Disclosure; SEQ ID NO 101; 126pp; English.  
 XX  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryopoietic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a peptide-linker compound.  
 XX  
 SQ Sequence 44 AA;  
 Query Match 57.4%; Score 62; DB 7; Length 44;  
 Best Local Similarity 84.6%; Pred. No. 0.042;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLKGR 16  
 DB 30 DGPTLRQWLKGR 42  
 ID ADN59780 standard; protein; 46 AA.  
 AC ADN59780;  
 DT 01-JUL-2004 (first entry)  
 XX  
 DB Peptide-vehicle compound, seq id 132.  
 XX  
 KW Haemostatic; antihaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KW TWP; c-mpl receptor; platelet precursor; megakaryocyte;  
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;  
KW autoimmune haemolytic anaemia; Hughes' syndrome;  
KW lupoid thrombocytopaenia.  
XX  
XX Unidentified.  
OS  
PN WO2003031589-A2.  
XX  
XX 17-APR-2003.  
PD  
XX 11-OCT-2002; 2002WO-US032552.  
PF  
XX 11-OCT-2001; 2001US-0328666P.  
PR 10-OCT-2002; 2002US-00269806.  
XX  
XX (AMGE-) AMGEN INC.  
PA  
XX Min H, Sitney KC, Hartley C;  
PI  
XX WPI; 2003-403101/38.  
DR  
XX  
XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
PT which stimulate the production of platelets and/or the production of  
PT platelet precursors, useful for treating thrombocytopaenia.  
XX  
XX  
XX Disclosure; SEQ ID NO 132; 126pp; English.  
XX  
XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
CC platelets and/or the production of platelet precursors, is new. Further  
CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
CC and a pharmaceutical composition comprising (II) and a carrier. The  
CC pharmaceutical composition of the invention is useful for treating  
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or  
CC platelets in a patient. The TMP of the invention is useful for treating  
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,  
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,  
CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid  
CC thrombocytopaenia. The TMP of the invention is also useful for  
CC maintaining the viability or storage life of platelets and/or  
CC megakaryocytes and its derived cells. The compounds demonstrate an  
CC improved ability to bind to and/or trigger transmembrane signal through,  
CC i.e. activating, the mpl receptor the compounds have superior  
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
CC platelet precursors. Further, certain of the compounds also exhibit  
CC superior therapeutic properties, such as improved plasma half-life,  
CC biological activity and in vivo circulation time. The current sequence  
CC represents a peptide-vehicle compound of the invention.  
XX  
XX  
SQ Sequence 46 AA;  
XX  
XX  
XX Query Match 57.4%; Score 62; DB 7; Length 46;  
XX Best Local Similarity 84.6%; Pred. No. 0.044;  
XX Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 DGPTLRQWLEGR 16  
DB 9 DGPTLRQWLEGR 21

XX  
XX Haemostatic; antihaemic; immunosuppressive; platelet;  
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
KW TWP; c-mpl receptor; platelet precursor; megakaryocyte;  
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;  
KW autoimmune haemolytic anaemia; Hughes' syndrome;  
KW lupoid thrombocytopaenia.  
XX  
XX Unidentified.  
OS  
PN WO2003031589-A2.  
XX  
XX 17-APR-2003.  
PD  
XX 11-OCT-2002; 2002WO-US032552.  
PF  
XX 11-OCT-2001; 2001US-0328666P.  
PR 10-OCT-2002; 2002US-00269806.  
XX  
XX (AMGE-) AMGEN INC.  
PA  
XX Min H, Sitney KC, Hartley C;  
PI  
XX WPI; 2003-403101/38.  
DR  
XX  
XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
PT which stimulate the production of platelets and/or the production of  
PT platelet precursors, useful for treating thrombocytopaenia.  
XX  
XX  
XX Disclosure; SEQ ID NO 138; 126pp; English.  
XX  
XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
CC platelets and/or the production of platelet precursors, is new. Further  
CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
CC and a pharmaceutical composition comprising (II) and a carrier. The  
CC pharmaceutical composition of the invention is useful for treating  
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or  
CC platelets in a patient. The TMP of the invention is useful for treating  
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,  
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,  
CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid  
CC thrombocytopaenia. The TMP of the invention is also useful for  
CC maintaining the viability or storage life of platelets and/or  
CC megakaryocytes and its derived cells. The compounds demonstrate an  
CC improved ability to bind to and/or trigger transmembrane signal through,  
CC i.e. activating, the mpl receptor the compounds have superior  
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
CC vitro, the production of platelets and/or megakaryocytopoietic activity,  
CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
CC platelet precursors. Further, certain of the compounds also exhibit  
CC superior therapeutic properties, such as improved plasma half-life,  
CC biological activity and in vivo circulation time. The current sequence  
CC represents a peptide-vehicle compound of the invention.  
XX  
XX  
SQ Sequence 46 AA;  
XX  
XX  
XX Query Match 57.4%; Score 62; DB 7; Length 46;  
XX Best Local Similarity 84.6%; Pred. No. 0.044;  
XX Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 DGPTLRQWLEGR 16  
DB 4 DGPTLRQWLEGR 16

RESULT 26  
ID ADN59786 standard; protein; 46 AA.  
XX  
XX ADN59786;  
AC  
XX 01-JUN-2004 (first entry)  
DT  
XX  
DE Peptide-vehicle compound, seq id 138.

RESULT 27  
ID ADN59667 standard; peptide; 18 AA.  
XX  
XX ADN59667;  
AC  
XX

DT 01-JUL-2004 (first entry)  
 XX Thrombopoietin mimetic peptide (TMP16), seq id 16.  
 DE  
 XX Haemostatic; antianaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KW autoimmune haemolytic anaemia; Hughes's syndrome;  
 KW lupoid thrombocytopenia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003031589-A2.  
 XX  
 PD 17-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032552.  
 XX  
 PR 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Min H, Sitney KC, Hartley C;  
 XX  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 XX  
 XX Claim 6; SEQ ID NO 16; 126bp; English.  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a preferred TMP of the invention.  
 XX  
 SQ Sequence 18 AA;  
 Query Match 55.6%; Score 60; DB 7; Length 18;  
 Best Local Similarity 71.4%; Pred. NO. 0.032;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Oy 3 ADGPTLRMLEGR 16  
 Db 3 AEGPTLRMLEGR 16  
 RESULT 28  
 ADNS9834  
 ID ADNS9834 standard; peptide; 22 AA.

XX ADNS9834;  
 AC  
 XX 01-JUL-2004 (first entry)  
 DT  
 XX  
 DE TMP peptide TMP16.  
 XX  
 KW Haemostatic; antianaemic; immunosuppressive; platelet;  
 KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
 KW autoimmune haemolytic anaemia; Hughes's syndrome;  
 KW lupoid thrombocytopenia; linker.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003031589-A2.  
 XX  
 PD 17-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032552.  
 XX  
 PR 11-OCT-2001; 2001US-0328666P.  
 PR 10-OCT-2002; 2002US-00269806.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Min H, Sitney KC, Hartley C;  
 XX  
 DR WPI; 2003-403101/38.  
 XX  
 PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
 PT which stimulate the production of platelets and/or the production of  
 PT platelet precursors, useful for treating thrombocytopenia.  
 XX  
 XX Example 6; Page 83; 126bp; English.  
 CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
 CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a TMP peptide of the invention to which a two amino acid "cap"  
 CC has been added to the carboxy terminal to increase peptide affinity.  
 XX  
 SQ Sequence 22 AA;  
 Query Match 55.6%; Score 60; DB 7; Length 22;  
 Best Local Similarity 71.4%; Pred. NO. 0.04;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Oy 3 ADGPTLRMLEGR 16  
 Db 5 AEGPTLRMLEGR 18



RESULT 29  
ADN59716  
ID ADN59716 standard; peptide; 25 AA.  
XX  
AC ADN59716;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Thrombopoietin mimetic peptide TMP16, seq id 65.  
XX  
KM Haemostatic; antihaemic; immunosuppressive; platelet;  
XX transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
XX TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
XX thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
XX autoimmune haemolytic anaemia; Hughes's syndrome;  
XX lupoid thrombocytopenia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003031589-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032552.  
XX  
XX 11-OCT-2001; 2001US-0328666P.  
XX  
PR 10-OCT-2002; 2002US-00269806.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Min H, Sitney KC, Hartley C;  
XX  
DR WPI; 2003-403101/38.  
XX  
DR N-PSDB; ADN59715.  
XX  
PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
XX which stimulate the production of platelets and/or the production of  
XX platelet precursors, useful for treating thrombocytopenia.  
XX  
PS Disclosure; SEQ ID NO 65; 126bp; English.  
XX  
XX  
XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
XX binds to the c-mpl (mpl) receptor, and which stimulates the production of  
XX platelets and/or the production of platelet precursors, is new. Further  
XX disclosed is a composition of matter (II) that binds to an mpl receptor,  
XX and a pharmaceutical composition comprising (II) and a carrier. The  
XX pharmaceutical composition of the invention is useful for treating  
XX thrombocytopenia in an animal, and for increasing megakaryocytes or  
XX platelets in a patient. The TMP of the invention is useful for treating  
XX conditions involving a megakaryocyte and/or platelet deficiency, e.g.  
XX disease conditions involving thrombocytopenia such as aplastic anaemia,  
XX autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
XX autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
XX thrombocytopenia. The TMP of the invention is also useful for  
XX maintaining the viability or storage life of platelets and/or  
XX megakaryocytes and its derived cells. The compounds demonstrate an  
XX improved ability to bind to and/or trigger transmembrane signal through,  
XX i.e. activating, the mpl receptor the compounds have superior  
XX thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
XX vitro, the production of platelets and/or megakaryocytopoietic activity,  
XX i.e. the ability to stimulate, in vivo and in vitro, the production of  
XX platelet precursors. Further, certain of the compounds also exhibit  
XX superior therapeutic properties, such as improved plasma half-life,  
XX biological activity and in vivo circulation time. The current sequence  
XX represents a TMP fragment.  
XX  
SQ Sequence 25 AA;

Query Match 55.6%; Score 60; DB 7; Length 25;  
Best Local Similarity 71.4%; Pred. No. 0.046;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 ADGPTLRQWLEGR 16  
|:|||||:|:|:  
Db 6 AEGPTLRQWLEGR 19

RESULT 30  
AA96527  
ID AA96527 standard; peptide; 34 AA.  
XX  
AC AA96527;  
XX  
DT 04-SEP-2000 (first entry)  
XX  
DE Thrombopoietin mimetic peptide compound 8.  
XX  
XX Thrombopoietin; mimetic; TMP; TPO; platelet; megakaryocyte; production;  
XX anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;  
XX immunosuppressive; anti-inflammatory; linker.  
XX  
OS Synthetic.  
XX  
XX  
XX Key Location/Qualifiers  
XX Modified-site 1  
XX Peptide /note= "optionally linked to an Fc molecule"  
XX Peptide 3..16  
XX Peptide /label= TMP\_1  
XX Peptide 17..20  
XX Peptide /label= linker  
XX Peptide 21..34  
XX Peptide /label= TMP\_2  
XX  
PN WO200024770-A2.  
XX  
XX 04-MAY-2000.  
XX  
XX 22-OCT-1999; 99WO-US024834.  
XX  
XX 23-OCT-1998; 98US-0105348P.  
XX  
XX (AMGE-) AMGEN INC.  
XX  
PI Liu C, Feige U, Cheetham J;  
XX  
DR WPI; 2000-365108/31.  
XX  
XX Thrombopoietic peptides which activate mpl receptors and increase the  
XX production of platelets or platelet precursors, useful for treatment of  
XX diseases which involve thrombocytopenia.  
XX  
PS Claim 16; Page 64; 91bp; English.  
XX  
XX A compound which binds to an mpl receptor comprising a thrombopoietin  
XX mimetic peptide (TMP) dimer joined by a linker (TMP\_1-(L\_1) nTMP\_2), is  
XX new. TMP\_1 and TMP\_2 are amino acid sequences varying from at least 10 to  
XX 14 residues in length comprising X\_2-X\_1\_0, X\_2-X\_1\_1, X\_2-X\_1\_2, X\_2-  
XX X\_1\_3, X\_2-X\_1\_4, X\_1-X\_1\_0, X\_1-X\_1\_1, X\_1-X\_1\_2, X\_1-X\_1\_3 and X\_1-  
XX X\_1\_4. X\_1 = I, A, V, L, S or R; X\_2 = E, D, K or V; X\_3 = G or A; X\_4 =  
XX P; X\_5 = T or S; X\_6 = L, I, V, A or F; X\_7 = R or K; X\_8 = Q, N, or E;  
XX X\_9 = W, Y or F; X\_1\_0 = L, I, V, A, F, M, or K; X\_1\_1 = A, I, V, L, F,  
XX S, T, K, H, or E; X\_1\_2 = A, I, V, L, F, G, S, or Q; X\_1\_3 = R, K, T, V,  
XX N, Q or G; X\_1\_4 = A, I, V, L, F, T, R, E, or G; L\_1 = linker comprising  
XX 1 to 20 amino acids, and n = 0 or 1. The compounds bind to and activate  
XX the c-mpl receptor which mediates the activity of endogenous  
XX thrombopoietin. The TMPs are useful for increasing the production of  
XX platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which  
XX is useful for treatment of diseases which involve thrombocytopenia, e.g.  
XX aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency  
XX virus associated ITP, and systemic lupus erythematosus  
XX  
SQ Sequence 34 AA;

Query Match 55.6%; Score 60; DB 3; Length 34;  
Best Local Similarity 57.9%; Pred. No. 0.065;



Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GNADGPTLRQWLEGGRRPRX 19  
 1 GIGEGPTLRQWLAARAGPN 19

# RESULT 31

ADNS9818  
 ID ADNS9818 standard; peptide; 42 AA.

AC ADNS9818;

DT 01-JUN-2004 (first entry)

DE Peptide-linker compound, seq id 102.

XX Haemostatic; antihaemic; immunosuppressive; platelet;

KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KM TGF- $\beta$  receptor; platelet precursor; megakaryocyte;

KM thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KM lupoid thrombocytopenia; linker.

XX Unidentified.

XX WO2003031589-A2.

XX 17-APR-2003.

XX 11-OCT-2002; 2002WO-US032552.

XX 11-OCT-2001; 2001US-032866P.

XX 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

XX Min H, Stiney KC, Hartley C;

XX WPI; 2003-403101/38.

XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

PT which stimulate the production of platelets and/or the production of

PT platelet precursors, useful for treating thrombocytopenia.

XX Disclosure; SEQ ID NO 102; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that

CC binds to the c-mpl (mpl) receptor, and which stimulates the production of

CC platelets and/or the production of platelet precursors, is new. Further

CC disclosed is a composition of matter (II) that binds to an mpl receptor,

CC and a pharmaceutical composition comprising (II) and a carrier. The

CC pharmaceutical composition of the invention is useful for treating

CC thrombocytopenia in an animal, and for increasing megakaryocytes or

CC platelets in a patient. The TMP of the invention is useful for treating

CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.

CC disease conditions involving thrombocytopenia such as aplastic anaemia,

CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,

CC thrombocytopenia. The TMP of the invention is also useful for

CC maintaining the viability or storage life of platelets and/or

CC megakaryocytes and its derived cells. The compounds demonstrate an

CC improved ability to bind to and/or trigger transmembrane signal through,

CC i.e. activating, the mpl receptor the compounds have superior

CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in

CC vitro, the production of platelets and/or megakaryocytic activity,

CC i.e. the ability to stimulate, in vivo and in vitro, the production of

CC platelet precursors. Further, certain of the compounds also exhibit

CC superior therapeutic properties, such as improved plasma half-life,

CC biological activity and in vivo circulation time. The current sequence

CC represents a peptide-linker compound.

XX Sequence 42 AA;

Query Match 55.6%; Score 60; DB 7; Length 42;

Best Local Similarity 71.4%; Pred. No. 0.081; Mismatches 1; Indels 0; Gaps 0;

Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 ADGPTLRQWLEGGRR 16  
 3 AEGPTLRQWLEGGRR 16

DB 3 AEGPTLRQWLEGGRR 16

# RESULT 32

ABP51670  
 ID ABP51670 standard; peptide; 15 AA.

AC ABP51670;

DT 01-OCT-2002 (first entry)

DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:2.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

KM complementarity determining region; immunoglobulin; antihaemic;

KM haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

XX Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

XX 04-MAY-2001; 2001US-0288899P.

XX 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXON PHARM INC.

XX Bowdish KS, Barbae-Fredrickson S, Renshaw M;

XX WPI; 2002-566610/60.

XX A novel immunogen molecule comprising a region in which amino acid

PT residues corresponding to at least a portion of the complementary

PT determining region are replaced or fused with an erythropoietin or

PT thrombopoietin mimetic.

XX Claim 19; Page 6; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment

CC (I) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (I) has

CC antihaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of

CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or

CC therapeutics, in cell isolation strategies, and for treating patients

CC suffering from deficiency in cell populations caused by disease,

CC disorders or treatments related to the suppression of haematopoiesis.

CC AB073288 to AB073377 and ABP51669 to ABP51696 represent sequences used in

CC the exemplification of the present invention

SQ Sequence 15 AA;  
 Query Match 54.6%; Score 59; DB 5; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 0.038;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLEGRRP 17  
 :|||||||  
 Db 2 EGPTLRQWLARAP 15  
 :|||||||  
 RESULT 33  
 ADQ16585  
 ID ADQ16585 standard; peptide; 15 AA.  
 XX  
 AC ADQ16585;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:2.  
 XX  
 DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.  
 XX  
 OS Unidentified.  
 OS  
 PN WO2004050017-A2.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PD 17-NOV-2003; 2003WO-US036894.  
 PF  
 XX 02-DEC-2002; 2002US-00307724.  
 PR  
 XX (ALEX-) ALEXION PHARM INC.  
 PA  
 PI Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.  
 XX  
 DR New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 PT  
 XX  
 PS Disclosure; SEQ ID NO 2; 107pp; English.  
 PS  
 CC The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide.  
 CC  
 CC  
 SQ Sequence 15 AA;  
 Query Match 54.6%; Score 59; DB 8; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 0.038;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLEGRRP 17  
 :|||||||  
 Db 2 EGPTLRQWLARAP 15  
 :|||||||  
 RESULT 34  
 ABP51689  
 ID ABP51689 standard; peptide; 18 AA.  
 XX

AC ABP51689;  
 XX  
 DT 01-OCT-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:41.  
 XX  
 DE TPO, EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200246238-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PD 05-DEC-2001; 2001WO-US047656.  
 PF  
 XX 05-DEC-2000; 2000US-0251448P.  
 PR 04-MAY-2001; 2001US-0288899P.  
 PR 29-MAY-2001; 2001US-0294068P.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Bowdish KS, Barbas-Frederickson S, Renshaw M;  
 XX WPI; 2002-566610/60.  
 XX  
 DR N-PSDB; ABQ73367.  
 XX  
 DR A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementarity  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.  
 PT  
 XX  
 PS Claim 20; Fig 5; 113pp; English.  
 PS  
 CC The present invention describes an immunoglobulin molecule or its fragment  
 CC (1) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementarity determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (1) has  
 CC antianemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (1) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (1) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (1) is contacted with haematopoietic  
 CC stem cells or their progenitors. (1) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51699 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 CC  
 CC  
 SQ Sequence 18 AA;  
 Query Match 54.6%; Score 59; DB 5; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.046;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 DGPTLRQWLEGRRP 17  
 :|||||||  
 Db 4 EGPTLRQWLARAP 17  
 :|||||||  
 RESULT 35  
 ABP51688  
 ID ABP51688 standard; peptide; 18 AA.  
 XX

XX	ABP51688 standard; peptide; 18 AA.
AC	ABP51688;
XX	
DT	01-OCT-2002 (first entry)
XX	
DE	TPO mimetic peptide SEQ ID NO:39.
XX	
KW	TPO, EPO, thrombopoietin; erythropoietin; antibody; CDR region;
KM	complementarily determining region; immunoglobulin; antianemic;
XX	haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
OS	Homo sapiens.
XX	Synthetic.
PN	WO200246238-A2.
PD	13-JUN-2002.
PF	05-DEC-2001; 2001WO-US047656.
PR	05-DEC-2000; 2000US-0251448P.
PR	04-MAY-2001; 2001US-0288889P.
PR	29-MAY-2001; 2001US-0294068P.
PA	(ALEX-) ALEXION PHARM INC.
PI	
XX	Bowdich KS, Barbas-Frederickson S, Renshaw M;
XX	WPI; 2002-566610/60.
DR	N-PSDB; ABO73366.
XX	
PT	A novel immunogen molecule comprising a region in which amino acid
PT	residues corresponding to at least a portion of the complementary
PT	determining region are replaced or fused with an erythropoietin or
XX	thrombopoietin mimetic.
XX	
PS	Claim 20; Fig 5; 113pp; English.
XX	
CC	The present invention describes an immunoglobulin molecule or its fragment
CC	(I) comprising a region where amino acid residues corresponding to at
CC	least a portion of the complementary determining region (CDR) are
CC	replaced or fused with biologically active peptides e.g. a peptide
CC	mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC	that is flanked with proline at its carboxy terminus. (I) has
CC	antianemic, haemostatic and nephrotropic activities, and can be used as
CC	a stimulator of proliferation, differentiation and maturation of
CC	haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC	for stimulating proliferation, differentiation or growth of
CC	promegakaryocytes or megakaryocytes, where (I) is contacted with
CC	promegakaryocytes or megakaryocytes, which results in increased platelet
CC	production. (I) with a region where amino acid residues corresponding to
CC	a portion of CDR is replaced with an EPO mimetic, or which has one or
CC	more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC	production of red blood cells, where (I) is contacted with haematopoietic
CC	stem cells or their progenitors. (I) is useful for diagnostics or
CC	therapeutics, in cell isolation strategies, and for treating patients
CC	suffering from deficiency in cell populations caused by disease,
CC	diagnosed or treatments related to the suppression of haematopoiesis.
CC	ABO733288 to ABO73377 and ABP51669 to ABP51696 represent sequences used in
CC	the exemplification of the present invention
XX	
XX	Sequence 18 AA;
XX	
Query Match	54.6%; Score 59; DB 5; Length 18;
Best Local Similarity	71.4%; Pred. No. 0.046;
Matches	10; Conservative 1; Mismatches 3; Indels 0; Gaps 0.
QY	4 DGPTRQWLEGRAP 17
DB	4 EGPTRLQWLAARAP 17

Query Match	54.6%	Score 59	DB 5	Length 18
Best Local Similarity	71.4%	Pred. No. 0.046		
Matches	10	Conservative	1	Mismatches 3; Indels 0; Gaps 0

  

Query	4 DGPRLKRWLEGRRP	17
QY	4 DGPRLKRWLEGRRP	17
DB	4 DGPRLKRWLEGRRP	17

## RESULT 37

ABP51693  
ID ABP51693 standard; peptide; 18 AA.

AC ABP51693;

DT 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:49.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
KW complementarity determining region; immunoglobulin; antianaemic;  
KM haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.  
OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

DR WPI: 2002-566610/60.

DR N-PSDB; ABQ73371.

PT A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
CC (1) comprising a region where amino acid residues corresponding to at  
CC least a portion of the complementary determining region (CDR) are  
CC replaced or fused with biologically active peptides e.g. a peptide  
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
CC that is flanked with proline at its carboxy terminus. (1) has  
CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
CC a stimulator of proliferation, differentiation and maturation of  
CC haematopoietic cells, and a stimulator of haematopoiesis. (1) is useful  
CC for stimulating proliferation, differentiation or growth of  
CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
CC promegakaryocytes or megakaryocytes, which results in increased platelet  
CC production. (1) with a region where amino acid residues corresponding to  
CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
CC production of red blood cells, where (1) is contacted with haematopoietic  
CC stem cells or their progenitors. (1) is useful for diagnostics or  
CC therapeutics, in cell isolation strategies, and for treating patients  
CC suffering from deficiency in cell populations caused by disease,  
CC disorders or treatments related to the suppression of haematopoiesis.  
CC ABQ73288 to ABQ73377 and ABP51693 to ABP51696 represent sequences used in  
CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 54.6%; Score 59; DB 5; Length 18;

Best Local Similarity 71.4%; Pred. No. 0.046;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DGPLTQWLEGRAP 17

DB 4 EGPLTQWLEGRAP 17

## RESULT 38

ABP51684  
ID ABP51684 standard; peptide; 18 AA.

AC ABP51684;

DT 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:31.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
KW complementarity determining region; immunoglobulin; antianaemic;  
KM haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.  
OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

DR WPI: 2002-566610/60.

DR N-PSDB; ABQ73362.

PT A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
CC (1) comprising a region where amino acid residues corresponding to at  
CC least a portion of the complementary determining region (CDR) are  
CC replaced or fused with biologically active peptides e.g. a peptide  
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
CC that is flanked with proline at its carboxy terminus. (1) has  
CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
CC a stimulator of proliferation, differentiation and maturation of  
CC haematopoietic cells, and a stimulator of haematopoiesis. (1) is useful  
CC for stimulating proliferation, differentiation or growth of  
CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
CC promegakaryocytes or megakaryocytes, which results in increased platelet  
CC production. (1) with a region where amino acid residues corresponding to  
CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
CC production of red blood cells, where (1) is contacted with haematopoietic  
CC stem cells or their progenitors. (1) is useful for diagnostics or  
CC therapeutics, in cell isolation strategies, and for treating patients  
CC suffering from deficiency in cell populations caused by disease,  
CC disorders or treatments related to the suppression of haematopoiesis.  
CC ABQ73288 to ABQ73377 and ABP51693 to ABP51696 represent sequences used in  
CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 54.6%; Score 59; DB 5; Length 18;

Best Local Similarity 71.4%; Pred. No. 0.046;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DGPTLRQWLEGRRP 17  
:|||||  
Db 4 EGPTLRQWLAAAP 17

## RESULT 39

ABP51691  
ID ABP51691 standard; peptide; 18 AA.

AC ABP51691;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:45.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
KM complementarity determining region; immunoglobulin; antianaemic;  
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

XX PA Bowdish KS, Barbae-Frederickson S, Renshaw M;

PI WPI; 2002-566610/60.

DR N-PSDB; ABQ73369.

XX A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.

XX PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment

CC (I) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (I) has

CC antianaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of

CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or

CC therapeutics, in cell isolation strategies, and for treating patients

CC suffering from deficiency in cell populations caused by disease.

CC disorders or treatments related to the suppression of haematopoiesis.

CC ABQ73288 to ABQ73377 and ABP51691 to ABP51696 represent sequences used in

CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 54.6%; Score 59; DB 5; Length 18;

Best Local Similarity 71.4%; Pred. No. 0.046;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DGPTLRQWLEGRRP 17  
:|||||  
Db 4 EGPTLRQWLAAAP 17

## RESULT 40

ABP51690  
ID ABP51690 standard; peptide; 18 AA.

AC ABP51690;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:43.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
KM complementarity determining region; immunoglobulin; antianaemic;  
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

XX PA Bowdish KS, Barbae-Frederickson S, Renshaw M;

PI WPI; 2002-566610/60.

DR N-PSDB; ABQ73368.

XX A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.

XX PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment

CC (I) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (I) has

CC antianaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of

CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or

CC therapeutics, in cell isolation strategies, and for treating patients

CC suffering from deficiency in cell populations caused by disease.

CC disorders or treatments related to the suppression of haematopoiesis.

CC ABQ73288 to ABQ73377 and ABP51691 to ABP51696 represent sequences used in

CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 54.6%; Score 59; DB 5; Length 18;

Query Match 54.6%; Score 59; DB 5; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.046;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DGPTRLQWLEGRRP 17  
 :|||||||  
 Db 4 EGPTLRQWLAARAP 17

## RESULT 41

ABP51675  
 ID ABP51675 standard; peptide; 18 AA.

XX AC ABP51675;  
 XX

DT 01-OCT-2002 (first entry)

DE TPO mimetic antibody related peptide graft SEQ ID NO:66.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KM complementarity determining region; immunoglobulin; antianaemic;  
 KM haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX OS Homo sapiens.  
 OS Synthetic.

PN WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

PA (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;  
 XX WPI; 2002-566610/60.

PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

XX Example 4; Page 55; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (1) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (1) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (1) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (1) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (1) is contacted with haematopoietic  
 CC stem cells or their progenitors. (1) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 XX

SQ Sequence 18 AA:

Query Match 54.6%; Score 59; DB 5; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.046;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DGPTRLQWLEGRRP 17  
 :|||||||  
 Db 4 EGPTLRQWLAARAP 17

## RESULT 42

ADQ16611  
 ID ADQ16611 standard; peptide; 18 AA.

XX AC ADQ16611;  
 XX

DT 09-SEP-2004 (first entry)

DE TPO mimetic peptide with random flanking residues SEQ ID NO:31.

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KM immunotherapy; thrombocytopenia.

XX OS Unidentified.

PN WO2004050017-A2.

XX 17-JUN-2004.

XX 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.

DR N-PSDB; ADQ16612.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 1; SEQ ID NO 31; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.  
 XX

SQ Sequence 18 AA:

Query Match 54.6%; Score 59; DB 8; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.046;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DGPTRLQWLEGRRP 17  
 :|||||||  
 Db 4 EGPTLRQWLAARAP 17

## RESULT 43

ADQ16619

ID	ADQ16619	standard; peptide; 18 AA.
XX		
AC	ADQ16619;	
XX		
DT	09-SEP-2004	(first entry)
XX		
DE	TPO mimetic peptide with random flanking residues SEQ ID NO:39.	
XX		
KW	immunoglobulin; complementarily determining region; CDR; peptide mimetic;	
KV	erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;	
XX	immunochemistry; thrombocytopenia.	
OS	Unidentified.	
XX		
PN	WO2004050017-A2.	
XX		
PD	17-JUN-2004.	
XX		
PF	17-NOV-2003; 2003WO-US036894.	
XX		
PR	02-DEC-2002; 2002US-00307724.	
XX		
PA	(ALEX-) ALEXION PHARM INC.	
XX		
PI	Bowditch KS, Frederickson S, Renshaw M;	
XX		
DR	WPI; 2004-460973/43.	
XX	N-PESDB; ADQ16620.	
PT	New immunoglobulin molecule comprising a region, where two	
PP	complementarily determining regions (CDRs) are replaced with EPO mimetic	
PS	or a TPO mimetic, useful for treating thrombocytopenia.	
PS	Example 1; SEQ ID NO 39; 107pp; English.	
XX		
CC	The invention relates to a novel immunoglobulin molecule or its fragment	
CC	comprising a region where amino acid residues corresponding to at least a	
CC	portion of a two complementarily determining regions (CDRs) are replaced	
CC	with a peptide mimetic selected from an erythropoietin (EPO) mimetic and	
CC	a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the	
CC	invention has immunosuppressive activity, and may have a use in	
CC	immunochemistry. The immunoglobulin molecule is useful for diagnosing or	
CC	treating thrombocytopenia as a result of chemotherapy, bone marrow	
CC	transplantation, or chronic diseases such as idiopathic thrombocytopenia.	
CC	The present sequence represents a TPO mimetic peptide with flanking	
CC	residues.	
XX		
SQ	Sequence 18 AA;	
	Query Match	54.6%; Score 59; DB 8; Length 18;
	Best Local Similarity	71.4%; Pred. No. 0.046; Mismatches
	Matches 10; Conservative 1; Indels 0; Gaps 0.	
OY	4 DGPTLRQLWGRRP 17 :      4 EGPTLRQLWAARAP 17	
Db		
RESULT 44		
ID	ADQ16621	ADQ16621 standard; peptide; 18 AA.
AC	ADQ16621;	
XX		
DT	09-SEP-2004	(first entry)
XX		
DE	TPO mimetic peptide with random flanking residues SEQ ID NO:41.	
XX		
KW	immunoglobulin; complementarily determining region; CDR; peptide mimetic;	
KV	erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;	
XX	immunochemistry; thrombocytopenia.	
OS	Unidentified.	

```

XX      WO2004050017-A2.
XX
XX      17-JUN-2004.
XX
XX      17-NOV-2003; 2003WO-US036894.
XX
XX      02-DEC-2002; 2002US-00307724.
XX
XX      (ALEX-) ALEXION PHARM INC.
XX
XX      Bowdish KS, Frederickson S, Renshaw M;
XX
XX      WPI; 2004-460973/43.
XX
XX      N-PSDB; ADQ16622.
XX
XX      New immunoglobulin molecule comprising a region, where two
XX      complementarily determining regions (CDRs) are replaced with EPO mimetic
XX      or a TPO mimetic, useful for treating thrombocytopenia.
XX
XX      Example 1; SEQ ID NO 41; 107pp; English.
XX
XX      The invention relates to a novel immunoglobulin molecule or its fragment
XX      comprising a region where amino acid residues corresponding to at least a
XX      portion of a two complementarily determining regions (CDRs) are replaced
XX      with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
XX      a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
XX      invention has immunosuppressive activity, and may have a use in
XX      immunotherapy. The immunoglobulin molecule is useful for diagnosing or
XX      treating thrombocytopenia as a result of chemotherapy, bone marrow
XX      transplantation, or chronic diseases such as idiopathic thrombocytopenia.
XX      The present sequence represents a TPO mimetic peptide with flanking
XX      residues.
XX
XX      Sequence 18 AA:
XX
XX      Query Match          54.6%; Score 59; DB 8; Length 18;
XX      Best Local Similarity 71.4%; Pred. No. 0.046;
XX      Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX      QY      4 DGPTRQWLEGRRP 17
XX              :|||||
XX      DB      4 EGPTLRQWLAARAP 17
XX
XX
XX
XX
XX      RESULT 45
XX      ADQ16646
XX      ID      ADQ16646 standard; peptide; 18 AA.
XX
XX      AC      ADQ16646;
XX      XX
XX      DT      09-SEP-2004 (first entry)
XX
XX      DE      TPO mimetic peptide SEQ ID NO:65.
XX
XX      KW      immunoglobulin; complementarily determining region; CDR; peptide mimetic;
XX      KM      erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
XX      OS      immunotherapy; thrombocytopenia.
XX
XX      Unidentified.
XX
XX      WO2004050017-A2.
XX
XX      PD      17-JUN-2004.
XX
XX      PF      17-NOV-2003; 2003WO-US036894.
XX
XX      PR      02-DEC-2002; 2002US-00307724.
XX
XX      (ALEX-) ALEXION PHARM INC.
XX
XX      Bowdish KS, Frederickson S, Renshaw M;
XX
XX
XX

```

DR WPI; 2004-460973/43.  
DR N-PSDB; ADQ16645.

XX New immunoglobulin molecule comprising a region, where two  
PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 4; SEQ ID NO 66; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
CC comprising a region where amino acid residues corresponding to at least a  
CC portion of a two complementarity determining regions (CDRs) are replaced  
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
CC invention has immunosuppressive activity, and may have a use in  
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide of the invention.

XX SQ Sequence 18 AA;

Query Match 54.6%; Score 59; DB 8; Length 18;  
Best Local Similarity 71.4%; Pred. No. 0.046; 3; Indels 0; Gaps 0;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DGPTLRQWLEGRRP 17  
:|||||||  
Db 4 EGPTLRQWLAARAP 17

Search completed: September 1, 2005, 16:12:10  
Job time : 89.3453 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 70.6691 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-7

Perfect score: 108

Sequence: 1 GNADGPTLRQWLGRRPKN 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.5	55.1	297	2	Q7UG64 rhodopirell
2	59	54.6	319	2	Q9KRM5 streptomyc
3	53	49.1	191	2	Q7PFR0 streptomyc
4	52	48.1	252	2	Q8XPO9 ralsconia s
5	51	47.2	655	2	Q7XVU9 oryza sativ
6	50.5	46.8	168	2	Q9V492 drosophila
7	50.5	46.8	256	2	Q6LEMS bochrops ja
8	50	46.3	228	2	Q75LM1 oryza sativ
9	50	46.3	391	2	Q8T511 anophelies g
10	50	46.3	391	2	Q7PUW8 streptomyc
11	49	45.4	242	2	Q82AV3 streptomyc
12	49	45.4	271	2	Q8EUN8 mycoplasma
13	49	45.4	387	2	Q872W6 neurospora
14	49	45.4	531	2	Q92281 rhizobium m
15	48.5	44.9	531	2	Q61G39 photobacter
16	48	44.4	81	2	Q9NDL7 hydra magni
17	48	44.4	129	2	Q8DHX7 synecococc
18	48	44.4	139	2	Q9GSR9 hydra atren
19	48	44.4	161	2	Q6A743 propionibac
20	48	44.4	282	1	Y356 MYCEN
21	48	44.4	364	2	Q7XUX8 oryza sativ
22	48	44.4	1023	2	Q65WR1 oryza sativ
23	48	44.4	1189	2	Q7X521 oryza sativ
24	47.5	44.0	264	2	Q7UR10 rhodopirell
25	47	43.5	81	2	Q9NDJ5 clima formos
26	47	43.5	81	2	Q9NDL8 hydractinia
27	47	43.5	81	2	Q9NDL9 extreme sp.
28	47	43.5	133	2	Q87645 methylcoc
29	47	43.5	181	2	Q93KY0 streptomyc
30	47	43.5	352	1	IID12 PYRAE
31	47	43.5	452	2	Q6FXT5 candida gla

32	47	43.5	571	2	Q9P729 neurospora
33	47	43.5	986	1	Z445 MOUSE
34	47	43.5	1005	2	Q8MT33 drosophila
35	47	43.5	1005	2	Q9V512 drosophila
36	47	43.5	1171	2	Q9P3E2 neurospora
37	46.5	43.1	522	2	Q9LIW0 oryza sativ
38	46.5	43.1	690	1	EP42 HUMAN
39	46.5	43.1	103	2	Q950T6 caenorhabdi
40	46	42.6	157	2	Q9NMG3 homo sapien
41	46	42.6	243	2	Q9RKP9 streptomyc
42	46	42.6	249	2	Q8W2V0 oryza sativ
43	46	42.6	249	2	Q7G731 oryza sativ
44	46	42.6	357	2	Q9S1Q2 streptomyc
45	46	42.6	362	2	Q82ON3 streptomyc
46	46	42.6	377	2	Q82PK5 streptomyc
47	46	42.6	479	1	RPCL_AERPE
48	46	42.6	656	2	Q6AF49
49	46	42.6	1049	2	Q9XBP6
50	46	42.6	1183	2	Q8H044 oryza sativ
51	46	42.6	1546	2	Q73UP3
52	45.5	42.1	88	2	Q6Y815
53	45.5	42.1	236	2	Q69TG4
54	45.5	42.1	446	2	Q828C4 streptomyc
55	45	41.7	147	2	Q825A6 streptomyc
56	45	41.7	147	2	Q82FW7 streptomyc
57	45	41.7	151	2	Q826F1 streptomyc
58	45	41.7	272	2	Q7N6R2
59	45	41.7	280	1	Y356 MYCCE
60	45	41.7	290	2	Q8K417
61	45	41.7	325	2	Q85MB2
62	45	41.7	326	2	P95613 rhizobium g
63	45	41.7	340	2	Q95Y14 acariis suu
64	45	41.7	344	2	Q9HN36 halobacteri
65	45	41.7	345	2	Q6DA67 erwinia car
66	45	41.7	542	2	Q6N2A1 rhodopseudo
67	45	41.7	559	2	Q66DB0 yersinia ps
68	45	41.7	559	2	Q8ZH10 yersinia pe
69	45	41.7	589	2	Q8CZ27 yersinia pe
70	45	41.7	609	2	Q856X8 mycobacteri
71	45	41.7	636	2	Q656W4 oryza sativ
72	45	41.7	863	2	Q9ST50 zea mays (m
73	45	41.7	1247	2	Q9SD34 arabidopsis
74	45	41.7	1413	2	Q9NBD3 caenorhabdi
75	45	41.7	1493	2	Q9NED3 nitrosomona
76	44.5	41.2	152	2	Q82X11 xanthomonas
77	44.5	41.2	173	2	Q8PLV3 streptomyc
78	44.5	41.2	446	2	Q88053 streptomyc
79	44.5	41.2	600	2	Q93IU2 streptomyc
80	44.5	41.2	602	2	P72407 streptomyc
81	44.5	41.2	1061	2	Q9W699 fugu rubrip
82	44.5	41.2	1456	2	Q8NUS1 leptosphaer
83	44	40.7	125	2	Q84HM5 rhizobium s
84	44	40.7	127	2	Q86MB0 lytechinus
85	44	40.7	189	2	Q9HSE8 halobacteri
86	44	40.7	192	2	Q7U7Q3 synecococc
87	44	40.7	210	1	MDCG_PSEAE
88	44	40.7	259	2	Q837Y4 enterococcu
89	44	40.7	260	2	Q8DVQ5 streptomyc
90	44	40.7	263	2	Q9PT52 agkistrodon
91	44	40.7	270	2	Q8EP23 oceanobacil
92	44	40.7	273	2	Q66D70 yersinia ps
93	44	40.7	273	2	Q8ZGK5 yersinia pe
94	44	40.7	302	2	Q742B3 mycobacteri
95	44	40.7	303	2	Q8FSC5 corynebacte
96	44	40.7	310	2	Q9MB12 bacterioph
97	44	40.7	320	2	Q8VKZ7 mycobacteri
98	44	40.7	324	1	G3P1_GLORO
99	44	40.7	330	1	G3PC_LETWE
100	44	40.7	330	2	Q82MF0 leishmania

## ALIGNMENTS

```

RESULT 1
07Q0B4 PRELIMINARY; PRT; 297 AA.
AC 07Q0B4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Ordered locus names=RB6375;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Aumann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294144; CAD74759.1; -.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR003169; GYP.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00829; GYP. 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475f670f02c78b3b CRC64;

Query Match 55.1%; Score 59.5; DB 2; Length 297;
Best Local Similarity 68.8%; Pred. No. 0.77;
Matches 11; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 GNADPTLRQWL-EGR 15
Db 173 GPADGPTMKQWISGR 188

RESULT 2
09RKM5 PRELIMINARY; PRT; 319 AA.
AC 09RKM5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative Meir family transcriptional regulator.
GN ORFNames=SCD17.06c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornbly T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -|- STIMILARITY: Contains 1 HTM merr-type DNA-binding domain.
DR EMBL; AL393118; CAB56383.1; -.
DR GO; GO:0005622; C:intracellular; IEA.

```

```

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTM_Merr.
DR InterPro; IPR009061; Putativ_DNA_bind.
DR Pfam; PF00376; Merr; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTM_MERR. 1.
DR PROSITE; PS50937; HTM_MERR_2; 1.
KM Complete proteome; DNA-binding.
SQ SEQUENCE 319 AA; 34841 MW; 1f51905a8ba5365e CRC64;

Query Match 54.6%; Score 59; DB 2; Length 319;
Best Local Similarity 66.7%; Pred. No. 0.99;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNADPTLRQWLEGR 15
Db 255 GRPDGPELRWELAGR 269

RESULT 3
07PRF0 PRELIMINARY; PRT; 191 AA.
AC 07PRF0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000014364 (Fragment).
GN Name=ENSANG0000011875;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (Apr-2003) to the EMBL/Genbank/DBJ databases.
CC -|- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB0100851; EAA07337.2; -.
FT NON TER 1
SQ SEQUENCE 191 AA; 21826 MW; 3DB9B8C839FAFCB CRC64;

Query Match 49.1%; Score 53; DB 2; Length 191;
Best Local Similarity 58.8%; Pred. No. 4.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 NADGPTLRQWLEGRPK 18
Db 74 HAAGPTERRWLEKESPK 90

RESULT 4
08XPQ9 PRELIMINARY; PRT; 252 AA.
AC 08XPQ9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
GN Name=RS02135; Ordered locus names=RSp1579;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1100;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

```

RA Arlat M., Billault A., Brothier P., Camus J.C., Catolico L.,  
 RA Chandler M., Christine N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulier P., Thebaud P., Whalen M., Winkler P., Levy M.,  
 RA Weisenbach J., Boucher C.A.,  
 RT "genome sequence of the plant pathogen *Ralstonia solanacearum*,"  
 RL Nature 415:497-502(2002).  
 CC -1- SIMILARITY: Contains 1 HTH LuxR-type DNA-binding domain.  
 DR EMBL; AL646085; CAD18730.1; -.  
 DR HSSP; P11470; 1PSE.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.  
 DR Pfam; PF00196; Gers; 1.  
 DR PRINTS; PRO1590; HTHLUX.  
 DR Prodom; PD000307; HTH LuxR; 1.  
 DR SMART; SM00421; HTH\_LUXR; 1.  
 KM Complete proteome; DNA-binding; Plasmid; Transcription;  
 KW Transcription regulation.  
 SQ SEQUENCE 252 AA; 2766 MW; 483403EE326F7C2E CRC64;

Query Match 48.1%; Score 52; DB 2; Length 252;  
 Best Local Similarity 52.9%; Pred. No. 9.3;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GNADPTLRQWLEGRPP 17  
 DB 74 GGDTPIMRMLATRRP 90

RESULT 5  
 Q7XVU9 PRELIMINARY; PRT; 655 AA.  
 ID Q7XVU9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)  
 DE OSJNB0035B13.2 protein.  
 GN Name=OSJNB0035B13.2;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Erihatoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=39947;  
 OX RN  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12447439; DOI=10.1038/nature01183;  
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,  
 RA Wang Q., Zhang L., Lu Y., Mu Y., Zhang L.S., Yu Z., Fan D.,  
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
 RA Chen J., Kang H., Chen C., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
 RA Han B.;  
 RT "Sequence and analysis of rice chromosome 4,"  
 RL Nature 420:316-320(2002).  
 DR EMBL; AL662966; CAD0429.1; -.  
 DR Gramene; Q7XVU9; -.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
 DR InterPro; IPR000477; RYase.  
 DR Pfam; PF00078; RVT\_1; 2.  
 KM RNA-directed DNA polymerase; Transferase.  
 SQ SEQUENCE 655 AA; 74463 MW; 4F15BC16D1BC776 CRC64;

Query Match 47.2%; Score 51; DB 2; Length 655;  
 Best Local Similarity 50.0%; Pred. No. 37;

Matches 10; Conservative 4; Mismatches 2; Indels 4; Gaps 1;  
 QY 4 DEPTLR---QWLEGRPPK 19  
 DB 483 DGNTIRFESAMWDGGRPK 502

RESULT 6  
 Q9V492 PRELIMINARY; PRT; 168 AA.  
 ID Q9V492;  
 AC Q9V492;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE CG11077-PA (RES5125P).  
 GN ORENames=CG11077;  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 OX RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blaziel R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
 RA Burdick K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Doonan K., Doup L.E., Downes W., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertire S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemison U.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Slier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.W., Weisenbach J.,  
 RA Williams S.M., Woodaght, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang K., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers B.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila melanogaster* euchromatic genome sequence,"

```

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celisner S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
RA Beutencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RP [6]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RP [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Bokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe W., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celisner S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO03846; AAF59393.1; -
DR EMBL; AY071482; AAL49104.1; -
DR FLYBASE; FBgn003930; CG11077.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005778; C:peroxisomal membrane; IEA.
DR InterPro; IPR007248; Mpv17_PMP22.
DR Pfam; PF04117; Mpv17_PMP22; 1.
SQ SEQUENCE 168 AA; 19521 MW; 48E216A954E3D39 CRC64;

Query Match 46.8%; Score 50.5; DB 2; Length 168;
Best Local Similarity 64.7%; Pred. No. 10;
Matches 11; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 5 GPTLRQW---LEGRRPK 18
Db |||||:|||||
54 GPTLRWYHFLSRVPK 70

RESULT 7
Q6LEMS PRELIMINARY; PRT; 256 AA.
ID 06LEMS
AC 06LEMS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Bradykinin-potentiating peptides and C-type natriuretic peptide.
OS Bothrops jararaca (Jarataca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=97188443; PubMed=9037028; DOI=10.1073/pnas.94.4.1189;
RA Murayama N., Hayaishi M., Ohi H., Ferreira L., Hermann V., Saito H.,
RA Fujita Y., Higuchi S., Fernandes B., Yamane T., de Camargo A.;
RT "Cloning and sequence analysis of a Bothrops jararaca cDNA encoding a
precursor of seven bradykinin-potentiating peptides and a C-type
natriuretic peptide."
RL Proc. Natl. Acad. Sci. U.S.A. 94:1189-1193(1997).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the natriuretic peptide family.
DR EMBL; D85843; BAA12879.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR Prodom; PD005617; Natr_peptide; 1.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
KW Vasoactive.
SQ SEQUENCE 256 AA; 26814 MW; 85BDBA0A9520A45 CRC64;

Query Match 46.8%; Score 50.5; DB 2; Length 256;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10; Conservative 3; Mismatches 4; Indels 5; Gaps 1;

QY 1 GNADGP-----TLRWLEGRRP 17
Db |||||:|||||
86 GRAPGPPIPLVYQWAGRAP 107

RESULT 8
Q7SLML PRELIMINARY; PRT; 228 AA.
ID 07SLML
AC 07SLML;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNBA0047E24.16.
GN Name=OSUNBA0047E24.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitlin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrieh D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Ullrich T.T., Feldlyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC092556; AAR87260.1; -
KW Hypothetical protein
SQ SEQUENCE 228 AA; 25966 MW; 8B9E7D088D49F5F2 CRC64;

Query Match 46.3%; Score 50; DB 2; Length 228;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 4 DGPTLR---QLEGRPPK 19
Db |||||:|||||
99 DGNTARFWSAWIDGRPK 118

RESULT 9

```

```

08T511
ID 08T511      PRELIMINARY;      PRT;      391 AA.
AC 08T511;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Transcription factor.
GN Name=30E5.9;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RX MEDLINE=22056115; PubMed=12060762; DOI=10.1073/pnas.082235599;
RA Thomasova D., Ton L.O., Copley R.R., Zdobnov E.M., Wang X., Hong Y.S.,
RA Sim C., Bork P., Kafatos F.C., Collins F.H.;
RT "Comparative genomic analysis in the region of a major Plasmodium-
RT retractoriness locus of Anopheles gambiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:8179-8184(2002).
DR EMBL; AJ439353; CADD7931.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR008895; YL1.
DR Pfam; PF05764; YL1; 1.
SQ SEQUENCE 391 AA; 45110 MW; E3F5D6D396460A37 CRC64;

Query Match      46.3%; Score 50; DB 2; Length 391;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GNADPTLRWLGR 16
DB 337 GNTNDPVARLWWRX 352

RESULT 10
Q7PUM8      PRELIMINARY;      PRT;      391 AA.
ID 07PUM8;
AC 07PUM8;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000018140.
GN Name=ENSANG000000015651;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008987; EAA00898.2; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR008895; YL1.
DR Pfam; PF05764; YL1; 1.
SQ SEQUENCE 391 AA; 45110 MW; E3F5D6D396460A37 CRC64;

Query Match      46.3%; Score 50; DB 2; Length 391;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GNADPTLRWLGR 16
DB 337 GNTNDPVARLWWRX 352

```

```

DB 337 GNTNDPVARLWWRX 352

RESULT 11
082AY3      PRELIMINARY;      PRT;      242 AA.
ID 082AY3;
AC 082AY3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SAV5922;
OS Streptomyces avermitilis.
OC Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa U., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa U., Hanamoto A., Takahashi C.,
RA Kikuchi H., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kituchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AP005044; BAC73634.1; -.
RX Complete proteome.
SQ SEQUENCE 242 AA; 26447 MW; 9D435A8C94401C1C CRC64;

Query Match      45.4%; Score 49; DB 2; Length 242;
Best Local Similarity 72.7%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LRQWLEGRPRK 18
DB 15 ISCWLRGRPRK 25

RESULT 12
Q8EUN8      PRELIMINARY;      PRT;      271 AA.
ID 08EUN8;
AC 08EUN8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Predicted choline kinase.
GN OrderedLocustNames=KYP89480;
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=1246555; DOI=10.1093/nar/gkf667;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenti T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004174; BAC44735.1; -.
DR HSSP; Q22942; IMW1.
DR GO; GO:0016301; F:kinase activity; IEA.

```

DR InterPro: IPR002573; Choline kinase.  
 DR InterPro: IPR011009; Kinase like.  
 DR Pfam: PF01633; Choline kinase; 1.  
 KW Complete proteome; Kinase.

SO SEQUENCE 271 AA; 33289 MW; CFFADCC9C24247D CRC64;

Query Match 45.4%; Score 49; DB 2; Length 271;  
 Best Local Similarity 46.7%; Pred. No. 29;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 DEPTLRQWLEGRRP 18  
 DB 81 DGNAIKRWLEGNPK 95

## RESULT 13

Q872W6 PRELIMINARY; PRT; 387 AA.

AC Q872W6; 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

GN Hypothetical protein B2G14.020.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

NCBI\_TaxID=5141;

RP SEQUENCE FROM N.A.

RA Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,

RA Nyakatura G., Mewes H.W., Mannhaupt G.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

KW Hypothetical protein.

SO SEQUENCE 387 AA; 42885 MW; 8C6PE22F50D7599 CRC64;

Query Match 45.4%; Score 49; DB 2; Length 387;  
 Best Local Similarity 44.4%; Pred. No. 43;  
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 GNADGPTLRQWLEGRRP 18  
 DB 355 GNANGSRVHRMARGRROR 372

## RESULT 14

Q92281 PRELIMINARY; PRT; 531 AA.

AC Q92281; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

GN Hypothetical protein.

OR Names=SMa1131;

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

NCBI\_TaxID=382;

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RA MEDLINE=2136509; PubMed=11481432; DOI=10.1073/pnas.161294798;

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,

RA Barlow-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,

RA Gurial M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,

RA Kaiman S., Keating D.H., Palm C., Peck K.C., Surzycki R., Wells D.H.,

RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;

RT "Nucleotide sequence and predicted functions of the entire

RT Sinorhizobium meliloti pSymA megaplasmid.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).

DR EMBL: AE007250; AAK65269.1; -.

DR PIR: C95338; C95338.

DR InterPro: IPR001279; Bactlase-like.

DR InterPro: IPR011108; RMBL.

DR Pfam: PF00753; Lactamase\_B; 1.

DR Pfam: PF07521; RMBL; 1.

KW Complete proteome; Hypothetical protein; Plasmid.

SO SEQUENCE 531 AA; 58948 MW; 93BECTFB16752E1 CRC64;

## RESULT 15

Q61G99 PRELIMINARY; PRT; 551 AA.

AC Q61G99; 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

GN Putative 2,3-dihydroxybenzoate-AMP ligase.

OS Photobacterium profundum (Photobacterium sp. (strain SS9)).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Photobacterium.

NCBI\_TaxID=74109;

RP SEQUENCE FROM N.A.

RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,

RA Cestaro A., Malacrida G., Simonato B., Camata N., Bartlett D.,

RA Valle G.;

RT "Genome analysis of Photobacterium profundum reveals the complexity of

RT high pressure adaptations.";

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme

CC family.

DR EMBL: CR378680; CAG23681.1; -.

DR GO: GO:0003824; F: catalytic activity; IEA.

DR GO: GO:0008152; P: metabolism; IEA.

DR InterPro: IPR000873; AMP-bind.

DR Pfam: PF00501; AMP-binding; 1.

DR PROSITE: PS00455; AMP\_BINDING; UNKNOWN\_1.

KW Complete proteome.

SO SEQUENCE 551 AA; 60767 MW; B3BB9F7A4F4B4EAD CRC64;

Query Match 44.9%; Score 48.5; DB 2; Length 551;  
 Best Local Similarity 71.4%; Pred. No. 75;  
 Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 NADGPTLRQWLEGRRP 14  
 DB 147 NADPTLRQWLEGRRP 160

## RESULT 16

Q9NDL7 PRELIMINARY; PRT; 81 AA.

AC Q9NDL7; 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

GN Glyceroldehyde-3-phosphate dehydrogenase (Fragment).

OS Hydra magnipapillata (Hydra).

OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroids; Anthomedusae;

```

Ox Hydriidae; Hydra.
OC NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RA Mochizuki K.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis, first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
CC dehydrogenase family.
DR EMBL; AB044096; BAA96506.1; -.
DR HSSP; P06977; 1DC3.
DR GO; GO:0004366; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001173; GAP_dhdrogenase.
DR Pfam; PF02800; Gp_dh_C; 1.
DR Pfam; PF00044; Gp_dh_N; 1.
DR PRINTS; PR00078; G3PDHDKGNASE.
DR PROSITE; PS00071; GAPDH; 1.
KM Glycolysis; NAD; Oxidoreductase.
FT NON TER 1
FT NON TER 81
SQ SEQUENCE 81 AA; 8577 MW; D3733493FDIC50D3 CRC64;

Query Match 44.4%; Score 48; DB 2; Length 81;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 4 DGPTLRQWLEGR 15
Db 57 DGSFKWVRDGR 68

RESULT 17
OBDHX7
ID OBDHX7 PRELIMINARY; PRT; 129 AA.
AC OBDHX7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE T11816 protein.
CN OrderedLocustNames=t11816;
OS Synchococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shindo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Rep. 9:123-130(2002).
KW EMBL; AP005375; BAC039368.1; -.
KW Complete proteome.
SQ SEQUENCE 129 AA; 14644 MW; EBB44691E7DD1E12 CRC64;

Query Match 44.4%; Score 48; DB 2; Length 129;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GNADGPTLRQWL 12
Db 38 GRAAGATLRQWL 49

```

ID	ORGSP9	PRELIMINARY;	PRT;	139 AA.
AC	09GSP9.			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, last annotation update)			
DE	Glyceralddehyde-3-phosphate dehydrogenase (Fragment)			
OS	Hydra attenuata (Hydra) (Hydra vulgaris)			
OC	Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;			
OC	Hydridae; Hydra.			
OX	NCBI_TaxID=60877;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Soderstrom K., De Petrocellis L., Di Marzo V.;			
RU	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: D-glyceralddehyde 3-phosphate + phosphate +			
CC	NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.			
CC	-1- PATHWAY: Second phase of glycolysis; first step.			
CC	-1- SUBUNIT: Homotetramer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: Belongs to the glyceralddehyde-3-phosphate			
CC	dehydrogenase family.			
DR	EMBL; AF307863; AAG29828.1; -.			
DR	HSSG; P46406; 100X.			
DR	GO; G0:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. .; IEA.			
DR	GO; G0:0016491; F:oxidoreductase activity; IEA.			
DR	GO; G0:0006096; P:glycolysis; IEA.			
DR	InterPro; IPR000173; GAP_dhhydrogenase.			
DR	Pfam; PF02800; GP_dh_C; 1.			
DR	Pfam; PF00044; GP_dh_N; 1.			
DR	PRINTS; PR00078; G3PDHGRNASE.			
DR	PROSITE; PS00071; GAPdh; 1.			
KW	GLYCOLYSIS; NAD; Oxidoreductase.			
FT	NON_TER 1			
FT	NON_TER 139			
FT	NON_TER 1			
SQ	SEQUENCE 139 AA; 14689 MW; 55D8B0F533C1E32A CRC64;			
Query Match		44.4%;	Score 48;	DB 2;
Best Local Similarity		50.0%;	Pred. No. 20;	
Matches	6;	Conservative	5;	Mismatches 1;
				Indels 0;
				Gaps 0;
Qy	4 DGPTRQWLEGR 15			
Db	62 DGPSMKWEKRDGR 73			
RESULT 19				
ID	06A743	PRELIMINARY;	PRT;	161 AA.
AC	06A743;			
DT	25-OCT-2004 (TREMBLrel. 28, Created)			
DT	25-OCT-2004 (TREMBLrel. 28, last sequence update)			
DT	25-OCT-2004 (TREMBLrel. 28, last annotation update)			
DE	Conserved protein.			
GN	OrderedlocusNames=PPA1691;			
OS	Propionibacterium acnes.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Propionibacterinae; Propionibacteriaceae; Propionibacterium.			
OX	NCBI_TaxID=17477;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=KPA171202 / DSM 16379;			
RX	Pubmed=15286373; DOI=10.1126/science.1100330;			
RA	Bruggemann H., Henne A., Hofer F., Liesegang H., Wietzer A.,			
RA	Strittmatter A., Hujer S., Duerre P., Gottschalk G.;			
RT	"The complete genome sequence of Propionibacterium acnes, a commensal			
RT	of human skin."			
RL	Science 305:671-673(2004).			
DR	EMBL; AB017283; AAT83422.1; -.			
KW	Complete proteome.			
SQ	SEQUENCE 161 AA; 17762 MW; 52D0DF0CE1330F0E CRC64;			



Query Match 44.4%; Score 48; DB 2; Length 161;  
Best Local Similarity 52.9%; Pred. No. 24;  
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GNADPTLAWLEGRPP 17  
DB 84 GAEDVPTMDPTEGRVP 100

## RESULT 20

ID Y356\_MYCPN STANDARD; PRT; 282 AA.

AC P75246;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Hypothetical protein MG356 homolog (G12orf282b).  
GN OrderedLocNames=MpN532; ORFNames=Mp310;  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RA MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;  
RA Himmelfreid R., Hilbert H., Plagens H., Pirkil E., Li B.-C.,  
RA Hermann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: A800028; AAB5958.1; -  
DR PIR, S73636; S73636.  
DR InterPro, IPR011009; Kinase like.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 282 AA; 33295 MW; 5662F03B9E06A89A CRC64;

Query Match 44.4%; Score 48; DB 1; Length 282;  
Best Local Similarity 42.9%; Pred. No. 43;  
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 5 GPTLRQWLEGRPP 18  
DB 85 GNAIKMKIEGKPP 98

## RESULT 21

ID Q7XUX8 PRELIMINARY; PRT; 364 AA.

AC Q7XUX8;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE OSJNBa0027G07.11 protein.  
GN Name=OSJNBa0027G07.11;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PubMed=12447439; DOI=10.1038/nature01183;  
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,

RA Meng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,  
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
RA Han B.;  
RT "Sequence and analysis of rice chromosome 4.";  
RL Nature 420:316-320(2002).  
DR EMBL: AL662937; CAD40947.1; -  
DR Gramene: Q7XUX8; -  
SQ SEQUENCE 364 AA; 41484 MW; 734879D485FFD13 CRC64;

Query Match 44.4%; Score 48; DB 2; Length 364;  
Best Local Similarity 50.0%; Pred. No. 57;  
Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 4 DGPTLR----QWLEGRPPK 19  
DB 166 DGNTRFWESAWINGRRPKD 185

## RESULT 22

ID Q6SWR1 PRELIMINARY; PRT; 1023 AA.

AC Q6SWR1;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Putative polypeptide.  
GN Name=P0009H09.8;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
RA Chen Y.-L., Cheng C.-H., Chang C.-I., Han S.-Y., Hsiao S.-H.,  
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,  
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
RA Wu H.-P., Shaw J.-F.;  
RT "Oryza sativa PAC P0009H09 genomic sequence.";  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC144743; AAU44186.1; -  
KW Polypeptide.  
SQ SEQUENCE 1023 AA; 117011 MW; B2D7844463910046 CRC64;

Query Match 44.4%; Score 48; DB 2; Length 1023;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 4 DGPTLR----QWLEGRPPK 19  
DB 747 DGNTRFWDSAWINGRRPKD 766

## RESULT 23

ID Q7X521 PRELIMINARY; PRT; 1189 AA.

AC Q7X521;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE OSJNBa0006A01.14 protein.  
GN Name=OSJNBa0006A01.14;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.



```

OX NCBI_TaxID=39947;
RN
RP SEQUENCE FROM N.A.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL715179; CAD41559.2; -.
DR Gramene; 07X521; -.
DR GO; GO:0003373; F:RNA binding; IEA.
DR GO; GO:0003664; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT_1; 1.
DR KMA; KMA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1189 AA; 135959 MW; 441C5E4B0B8F2643 CRC64;

Query March 44.4%; Score 48; DB 2; Length 1189;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 4 DGPTLR-----QMLEGRPRPKN 19
    |||||
    913 DGNTRFWDMSAMINGRRPKD 932

RESULT 24
Q7UR10 PRELIMINARY; PRT; 264 AA.
ID 07UR10;
AC 07UR10;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RB5963;
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloebsch F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Lindqvist W., Gade D., Beck K., Borzym K., Heilmann K., Rabus R.,
RA Schleuter H., Aumann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294143; CAD74532.1; -.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 264 AA; 29816 MW; 932AF9B90DC59C19 CRC64;

Query March 44.0%; Score 47.5; DB 2; Length 264;
Best Local Similarity 43.5%; Pred. No. 48;
Matches 10; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

Qy 2 NADGP-----TLRVLGRPRPKN 19
    |||||
    205 NLDSPKTKAKRIRTLMBEHRPN 227

```

```

RESULT 25
Q9NDL5 PRELIMINARY; PRT; 81 AA.
ID Q9NDL5;
AC Q9NDL5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN Name=GAPDH;
OS Tima formosa.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Eumimidae; Tima.
OX NCBI_TaxID=128134;
RN [1]
RP SEQUENCE FROM N.A.
RA Mochizuki K.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde-3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
CC dehydrogenase family.
DR EMBL; AB044098; BA96508.1; -.
DR HSSP; P06977; IDC3.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001173; GAP_dhdrogenase.
DR Pfam; PF02800; Gp_dh_C; 1.
DR Pfam; PF00044; Gp_dh_N; 1.
DR PRINTS; PR00078; GAPDHGNASE.
DR PROSITE; PS00071; GAPDH; 1.
KM Glycolysis; NAD; Oxidoreductase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 81 AA; 8499 MW; 32D2FFCA3C6D1C23 CRC64;

Query March 43.5%; Score 47; DB 2; Length 81;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DGPTLRQWLEGR 15
    |||||
    57 DGPSMKKKRDR 68

RESULT 26
Q9NDL8 PRELIMINARY; PRT; 81 AA.
ID Q9NDL8;
AC Q9NDL8;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN Name=GAPDH;
OS Hydractinia echinata (Snail fur).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydractiniidae; Hydractinia.
OX NCBI_TaxID=35630;
RN [1]
RP SEQUENCE FROM N.A.
RA Mochizuki K.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde-3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
CC dehydrogenase family.

```



OY 2 NADGPTLRQWEG-----RRP 17  
 DB 147 DADDPVSVEWARGDPDRTRP 166

## RESULT 30

ID12\_PYRAE STANDARD; PRT; 352 AA.

AC Q82YF6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE (Isopentenyl)-diphosphate delta-isomerase (EC 5.3.3.2) (IPP isomerase)  
 DE (Isopentenyl)-pyrophosphate isomerase  
 GN Name=Ipi; OrderedLocustNames=PA080801;  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 CX NCBI\_TaxID=13773;  
 (1)

SEQUENCE FROM N.A.  
 RP STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RC MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 CC -1- FUNCTION: Catalyzes the 1,3-allylic rearrangement of the  
 homocallylic substrate isopentenyl (IPP) to its allylic isomer,  
 dimethylallyl diphosphate (DMAPP) (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate = dimethylallyl  
 diphosphate.  
 CC -1- COFACTOR: FMN and NADPH (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the IPP isomerase type 2 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AE009786; AAL63037.1; -  
 DR HSSP; P50740; 1POK.  
 DR HAMAP; MF\_00354; -, 1.  
 DR InterPro; IPR003009; FMN\_enzyme.  
 DR InterPro; IPR011179; IPdp\_isomerase.  
 DR PIRSF; PIRSF003314; IPP\_isomerase; 1.  
 KW Complete proteome; Flavoprotein; FMN; Isomerase;  
 KW Isoprene biosynthesis; NADP;  
 SO SEQUENCE 352 AA; 37966 MW; 684253886324C04 CRC64;

Query Match 43.5%; Score 47; DB 1; Length 352;  
 Best Local Similarity 66.7%; Pred. No. 79;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 CPTLRQWLEGR 16  
 DB 338 GPRLNWTEORR 349

## RESULT 31

ID Q6FXT5 PRELIMINARY; PRT; 452 AA.

AC Q6FXT5;  
 DT 05-JUL-2004 (TRENBLREL. 27, Created)  
 DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)  
 DE Similar to sp|P34221|Saccharomyces cerevisiae YBL056w PTC3.  
 GN ORFNames=CAGL0A043019;

OS Candida glabrata CBS138.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; microsporid Saccharomycetales; Candida.  
 CX NCBI\_TaxID=284593;  
 (1)

SEQUENCE FROM N.A.  
 RP STRAIN=CBS138;  
 RC Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casarogola S.,  
 RA Lafontaine I., de Montigny J., March C., Neveglise C., Talla E.,  
 RA Goffard N., Frangul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boissame A., Boyer J., Cartolico L., Confiantieri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,  
 RA Hancryste F., Hemeguin C., Jaumaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicoud J.M., Nikolaki M., Oztas S., Ozer-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straud M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts.";  
 RL Nature 430:35-44(2004).  
 DR EMBL; CR380947; CAG57847.1; -  
 DR GO; GO:0003824; P:catalytic activity; IEA.  
 DR InterPro; IPR002222; PP2C.  
 DR InterPro; IPR001932; PP2C-like.  
 DR Pfam; PF00481; PP2C; 1.  
 DR SMART; SM00332; PP2C; 1.  
 DR PROSITE; PS01032; PP2C; UNKNOWN 1.  
 SO SEQUENCE 452 AA; 48881 MW; EBF9E25621E3315D CRC64;

Query Match 43.5%; Score 47; DB 2; Length 452;  
 Best Local Similarity 56.2%; Pred. No. 1e+02;  
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 DGPTLRQWLEGRPKN 19  
 DB 296 DSETLEQWFERMRAXN 311

RESULT 32  
 ID Q9P729 PRELIMINARY; PRT; 571 AA.

AC Q9P729;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE Probable histone acetyltransferase.  
 GN Name=BD4.110;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetiales; Sordariaceae; Neurospora.  
 CX NCBI\_TaxID=5141;  
 (1)

SEQUENCE FROM N.A.  
 RP Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Manhaupt G.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0046961; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0005506; F:iron ion binding; IEA.  
 DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.

```

DR InterPro: IPR000793; ATPase_a/b_C.
DR InterPro: IPR006638; ELP3/MAB/NiB.
DR InterPro: IPR005910; ELP3_Ac_trans.
DR InterPro: IPR000182; GNSacetyl_trans.
DR InterPro: IPR007197; Radical_SAM.
DR Pfam: PF00583; Acetyltransferase_1.
DR Pfam: PF04055; Radical_SAM; 1.
DR SMART: SM00729; ELP3; 1.
DR TIGRFAMs: TIGR01211; ELP3; 1.
DR Transferrase.
KM SEQUENCE 571 AA; 64684 MW; 3A36C590C2BCAD6 CRC64;

Query Match 43.5%; Score 47; DB 2; Length 571;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DGPITROWLEGR 15
Db 555 DGEYMSKMDGR 566

RESULT 33
2445 MOUSE STANDARD; PRT; 986 AA.
ID _2445_MOUSE
AC QGR2V3; QGR216;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Zinc finger protein 445.
GN Name=Znf445;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J;
RC Zhou G., Wang J., Zhang Y.;
RT "Cloning of mouse zinc finger protein 445."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 345-986 FROM N.A.
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heileh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalek U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May be involved in transcriptional regulation.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: Belongs to the Krueppel C2H2-type zinc-finger protein
CC family.
CC -1- SIMILARITY: Contains 12 C2H2-type zinc fingers.
CC -1- SIMILARITY: Contains 1 KRAB domain.
CC -1- SIMILARITY: Contains 1 SCAN box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AY341877; AA024161.1; -.
DR EMBL: BC027167; AAH27167.1; -.
DR EMBL: BC034572; AAH34572.1; ALT_INIT.
DR MGI: MGI:2143340; AW610627.
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR003309; Treg_SCAN.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF01352; KRAB; 1.
DR Pfam: PF02023; SCAN; 1.
DR Pfam: PF00096; Zf-C2H2; 12.
DR ProDom: PD000003; Znf_C2H2; 6.
DR SMART: SM00349; KRAB; 1.
DR SMART: SM00431; LER; 1.
DR SMART: SM00355; Znf_C2H2; 12.
DR PROSITE: PS50805; KRAB; 1.
DR PROSITE: PS50804; SCAN_BOX; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 12.
KW DNA-binding; Metal-binding; Nuclear protein; Repeat;
KW Transcription regulation; Zinc-finger.
FT DOMAIN 52 134
FT STRAIN=C57BL/6J;
FT ZN_FING 470 492 C2H2-type 1.
FT ZN_FING 498 520 C2H2-type 1.
FT ZN_FING 553 575 C2H2-type 3.
FT ZN_FING 581 604 C2H2-type 4.
FT ZN_FING 634 656 C2H2-type 5.
FT ZN_FING 662 686 C2H2-type 6.
FT ZN_FING 718 740 C2H2-type 7.
FT ZN_FING 746 768 C2H2-type 8.
FT ZN_FING 796 818 C2H2-type 9.
FT ZN_FING 824 846 C2H2-type 10.
FT ZN_FING 933 955 C2H2-type 11.
FT ZN_FING 961 983 C2H2-type 12.
FT ZN_FING 993 1015 C2H2-type 13.
FT ZN_FING 1023 1045 C2H2-type 14.
FT ZN_FING 1053 1075 C2H2-type 15.
FT ZN_FING 1083 1105 C2H2-type 16.
FT ZN_FING 1113 1135 C2H2-type 17.
FT ZN_FING 1143 1165 C2H2-type 18.
FT ZN_FING 1173 1195 C2H2-type 19.
FT ZN_FING 1203 1225 C2H2-type 20.
FT ZN_FING 1233 1255 C2H2-type 21.
FT ZN_FING 1263 1285 C2H2-type 22.
FT ZN_FING 1293 1315 C2H2-type 23.
FT ZN_FING 1323 1345 C2H2-type 24.
FT ZN_FING 1353 1375 C2H2-type 25.
FT ZN_FING 1383 1405 C2H2-type 26.
FT ZN_FING 1413 1435 C2H2-type 27.
FT ZN_FING 1443 1465 C2H2-type 28.
FT ZN_FING 1473 1495 C2H2-type 29.
FT ZN_FING 1503 1525 C2H2-type 30.
FT ZN_FING 1533 1555 C2H2-type 31.
FT ZN_FING 1563 1585 C2H2-type 32.
FT ZN_FING 1593 1615 C2H2-type 33.
FT ZN_FING 1623 1645 C2H2-type 34.
FT ZN_FING 1653 1675 C2H2-type 35.
FT ZN_FING 1683 1705 C2H2-type 36.
FT ZN_FING 1713 1735 C2H2-type 37.
FT ZN_FING 1743 1765 C2H2-type 38.
FT ZN_FING 1773 1795 C2H2-type 39.
FT ZN_FING 1803 1825 C2H2-type 40.
FT ZN_FING 1833 1855 C2H2-type 41.
FT ZN_FING 1863 1885 C2H2-type 42.
FT ZN_FING 1893 1915 C2H2-type 43.
FT ZN_FING 1923 1945 C2H2-type 44.
FT ZN_FING 1953 1975 C2H2-type 45.
FT ZN_FING 1983 2005 C2H2-type 46.
FT ZN_FING 2013 2035 C2H2-type 47.
FT ZN_FING 2043 2065 C2H2-type 48.
FT ZN_FING 2073 2095 C2H2-type 49.
FT ZN_FING 2103 2125 C2H2-type 50.
FT ZN_FING 2133 2155 C2H2-type 51.
FT ZN_FING 2163 2185 C2H2-type 52.
FT ZN_FING 2193 2215 C2H2-type 53.
FT ZN_FING 2223 2245 C2H2-type 54.
FT ZN_FING 2253 2275 C2H2-type 55.
FT ZN_FING 2283 2305 C2H2-type 56.
FT ZN_FING 2313 2335 C2H2-type 57.
FT ZN_FING 2343 2365 C2H2-type 58.
FT ZN_FING 2373 2395 C2H2-type 59.
FT ZN_FING 2403 2425 C2H2-type 60.
FT ZN_FING 2433 2455 C2H2-type 61.
FT ZN_FING 2463 2485 C2H2-type 62.
FT ZN_FING 2493 2515 C2H2-type 63.
FT ZN_FING 2523 2545 C2H2-type 64.
FT ZN_FING 2553 2575 C2H2-type 65.
FT ZN_FING 2583 2605 C2H2-type 66.
FT ZN_FING 2613 2635 C2H2-type 67.
FT ZN_FING 2643 2665 C2H2-type 68.
FT ZN_FING 2673 2695 C2H2-type 69.
FT ZN_FING 2703 2725 C2H2-type 70.
FT ZN_FING 2733 2755 C2H2-type 71.
FT ZN_FING 2763 2785 C2H2-type 72.
FT ZN_FING 2793 2815 C2H2-type 73.
FT ZN_FING 2823 2845 C2H2-type 74.
FT ZN_FING 2853 2875 C2H2-type 75.
FT ZN_FING 2883 2905 C2H2-type 76.
FT ZN_FING 2913 2935 C2H2-type 77.
FT ZN_FING 2943 2965 C2H2-type 78.
FT ZN_FING 2973 2995 C2H2-type 79.
FT ZN_FING 3003 3025 C2H2-type 80.
FT ZN_FING 3033 3055 C2H2-type 81.
FT ZN_FING 3063 3085 C2H2-type 82.
FT ZN_FING 3093 3115 C2H2-type 83.
FT ZN_FING 3123 3145 C2H2-type 84.
FT ZN_FING 3153 3175 C2H2-type 85.
FT ZN_FING 3183 3205 C2H2-type 86.
FT ZN_FING 3213 3235 C2H2-type 87.
FT ZN_FING 3243 3265 C2H2-type 88.
FT ZN_FING 3273 3295 C2H2-type 89.
FT ZN_FING 3303 3325 C2H2-type 90.
FT ZN_FING 3333 3355 C2H2-type 91.
FT ZN_FING 3363 3385 C2H2-type 92.
FT ZN_FING 3393 3415 C2H2-type 93.
FT ZN_FING 3423 3445 C2H2-type 94.
FT ZN_FING 3453 3475 C2H2-type 95.
FT ZN_FING 3483 3505 C2H2-type 96.
FT ZN_FING 3513 3535 C2H2-type 97.
FT ZN_FING 3543 3565 C2H2-type 98.
FT ZN_FING 3573 3595 C2H2-type 99.
FT ZN_FING 3603 3625 C2H2-type 100.
FT ZN_FING 3633 3655 C2H2-type 101.
FT ZN_FING 3663 3685 C2H2-type 102.
FT ZN_FING 3693 3715 C2H2-type 103.
FT ZN_FING 3723 3745 C2H2-type 104.
FT ZN_FING 3753 3775 C2H2-type 105.
FT ZN_FING 3783 3805 C2H2-type 106.
FT ZN_FING 3813 3835 C2H2-type 107.
FT ZN_FING 3843 3865 C2H2-type 108.
FT ZN_FING 3873 3895 C2H2-type 109.
FT ZN_FING 3903 3925 C2H2-type 110.
FT ZN_FING 3933 3955 C2H2-type 111.
FT ZN_FING 3963 3985 C2H2-type 112.
FT ZN_FING 3993 4015 C2H2-type 113.
FT ZN_FING 4023 4045 C2H2-type 114.
FT ZN_FING 4053 4075 C2H2-type 115.
FT ZN_FING 4083 4105 C2H2-type 116.
FT ZN_FING 4113 4135 C2H2-type 117.
FT ZN_FING 4143 4165 C2H2-type 118.
FT ZN_FING 4173 4195 C2H2-type 119.
FT ZN_FING 4203 4225 C2H2-type 120.
FT ZN_FING 4233 4255 C2H2-type 121.
FT ZN_FING 4263 4285 C2H2-type 122.
FT ZN_FING 4293 4315 C2H2-type 123.
FT ZN_FING 4323 4345 C2H2-type 124.
FT ZN_FING 4353 4375 C2H2-type 125.
FT ZN_FING 4383 4405 C2H2-type 126.
FT ZN_FING 4413 4435 C2H2-type 127.
FT ZN_FING 4443 4465 C2H2-type 128.
FT ZN_FING 4473 4495 C2H2-type 129.
FT ZN_FING 4503 4525 C2H2-type 130.
FT ZN_FING 4533 4555 C2H2-type 131.
FT ZN_FING 4563 4585 C2H2-type 132.
FT ZN_FING 4593 4615 C2H2-type 133.
FT ZN_FING 4623 4645 C2H2-type 134.
FT ZN_FING 4653 4675 C2H2-type 135.
FT ZN_FING 4683 4705 C2H2-type 136.
FT ZN_FING 4713 4735 C2H2-type 137.
FT ZN_FING 4743 4765 C2H2-type 138.
FT ZN_FING 4773 4795 C2H2-type 139.
FT ZN_FING 4803 4825 C2H2-type 140.
FT ZN_FING 4833 4855 C2H2-type 141.
FT ZN_FING 4863 4885 C2H2-type 142.
FT ZN_FING 4893 4915 C2H2-type 143.
FT ZN_FING 4923 4945 C2H2-type 144.
FT ZN_FING 4953 4975 C2H2-type 145.
FT ZN_FING 4983 5005 C2H2-type 146.
FT ZN_FING 5013 5035 C2H2-type 147.
FT ZN_FING 5043 5065 C2H2-type 148.
FT ZN_FING 5073 5095 C2H2-type 149.
FT ZN_FING 5103 5125 C2H2-type 150.
FT ZN_FING 5133 5155 C2H2-type 151.
FT ZN_FING 5163 5185 C2H2-type 152.
FT ZN_FING 5193 5215 C2H2-type 153.
FT ZN_FING 5223 5245 C2H2-type 154.
FT ZN_FING 5253 5275 C2H2-type 155.
FT ZN_FING 5283 5305 C2H2-type 156.
FT ZN_FING 5313 5335 C2H2-type 157.
FT ZN_FING 5343 5365 C2H2-type 158.
FT ZN_FING 5373 5395 C2H2-type 159.
FT ZN_FING 5403 5425 C2H2-type 160.
FT ZN_FING 5433 5455 C2H2-type 161.
FT ZN_FING 5463 5485 C2H2-type 162.
FT ZN_FING 5493 5515 C2H2-type 163.
FT ZN_FING 5523 5545 C2H2-type 164.
FT ZN_FING 5553 5575 C2H2-type 165.
FT ZN_FING 5583 5605 C2H2-type 166.
FT ZN_FING 5613 5635 C2H2-type 167.
FT ZN_FING 5643 5665 C2H2-type 168.
FT ZN_FING 5673 5695 C2H2-type 169.
FT ZN_FING 5703 5725 C2H2-type 170.
FT ZN_FING 5733 5755 C2H2-type 171.
FT ZN_FING 5763 5785 C2H2-type 172.
FT ZN_FING 5793 5815 C2H2-type 173.
FT ZN_FING 5823 5845 C2H2-type 174.
FT ZN_FING 5853 5875 C2H2-type 175.
FT ZN_FING 5883 5905 C2H2-type 176.
FT ZN_FING 5913 5935 C2H2-type 177.
FT ZN_FING 5943 5965 C2H2-type 178.
FT ZN_FING 5973 5995 C2H2-type 179.
FT ZN_FING 6003 6025 C2H2-type 180.
FT ZN_FING 6033 6055 C2H2-type 181.
FT ZN_FING 6063 6085 C2H2-type 182.
FT ZN_FING 6093 6115 C2H2-type 183.
FT ZN_FING 6123 6145 C2H2-type 184.
FT ZN_FING 6153 6175 C2H2-type 185.
FT ZN_FING 6183 6205 C2H2-type 186.
FT ZN_FING 6213 6235 C2H2-type 187.
FT ZN_FING 6243 6265 C2H2-type 188.
FT ZN_FING 6273 6295 C2H2-type 189.
FT ZN_FING 6303 6325 C2H2-type 190.
FT ZN_FING 6333 6355 C2H2-type 191.
FT ZN_FING 6363 6385 C2H2-type 192.
FT ZN_FING 6393 6415 C2H2-type 193.
FT ZN_FING 6423 6445 C2H2-type 194.
FT ZN_FING 6453 6475 C2H2-type 195.
FT ZN_FING 6483 6505 C2H2-type 196.
FT ZN_FING 6513 6535 C2H2-type 197.
FT ZN_FING 6543 6565 C2H2-type 198.
FT ZN_FING 6573 6595 C2H2-type 199.
FT ZN_FING 6603 6625 C2H2-type 200.
FT ZN_FING 6633 6655 C2H2-type 201.
FT ZN_FING 6663 6685 C2H2-type 202.
FT ZN_FING 6693 6715 C2H2-type 203.
FT ZN_FING 6723 6745 C2H2-type 204.
FT ZN_FING 6753 6775 C2H2-type 205.
FT ZN_FING 6783 6805 C2H2-type 206.
FT ZN_FING 6813 6835 C2H2-type 207.
FT ZN_FING 6843 6865 C2H2-type 208.
FT ZN_FING 6873 6895 C2H2-type 209.
FT ZN_FING 6903 6925 C2H2-type 210.
FT ZN_FING 6933 6955 C2H2-type 211.
FT ZN_FING 6963 6985 C2H2-type 212.
FT ZN_FING 6993 7015 C2H2-type 213.
FT ZN_FING 7023 7045 C2H2-type 214.
FT ZN_FING 7053 7075 C2H2-type 215.
FT ZN_FING 7083 7105 C2H2-type 216.
FT ZN_FING 7113 7135 C2H2-type 217.
FT ZN_FING 7143 7165 C2H2-type 218.
FT ZN_FING 7173 7195 C2H2-type 219.
FT ZN_FING 7203 7225 C2H2-type 220.
FT ZN_FING 7233 7255 C2H2-type 221.
FT ZN_FING 7263 7285 C2H2-type 222.
FT ZN_FING 7293 7315 C2H2-type 223.
FT ZN_FING 7323 7345 C2H2-type 224.
FT ZN_FING 7353 7375 C2H2-type 225.
FT ZN_FING 7383 7405 C2H2-type 226.
FT ZN_FING 7413 7435 C2H2-type 227.
FT ZN_FING 7443 7465 C2H2-type 228.
FT ZN_FING 7473 7495 C2H2-type 229.
FT ZN_FING 7503 7525 C2H2-type 230.
FT ZN_FING 7533 7555 C2H2-type 231.
FT ZN_FING 7563 7585 C2H2-type 232.
FT ZN_FING 7593 7615 C2H2-type 233.
FT ZN_FING 7623 7645 C2H2-type 234.
FT ZN_FING 7653 7675 C2H2-type 235.
FT ZN_FING 7683 7705 C2H2-type 236.
FT ZN_FING 7713 7735 C2H2-type 237.
FT ZN_FING 7743 7765 C2H2-type 238.
FT ZN_FING 7773 7795 C2H2-type 239.
FT ZN_FING 7803 7825 C2H2-type 240.
FT ZN_FING 7833 7855 C2H2-type 241.
FT ZN_FING 7863 7885 C2H2-type 242.
FT ZN_FING 7893 7915 C2H2-type 243.
FT ZN_FING 7923 7945 C2H2-type 244.
FT ZN_FING 7953 7975 C2H2-type 245.
FT ZN_FING 7983 8005 C2H2-type 246.
FT ZN_FING 8013 8035 C2H2-type 247.
FT ZN_FING 8043 8065 C2H2-type 248.
FT ZN_FING 8073 8095 C2H2-type 249.
FT ZN_FING 8103 8125 C2H2-type 250.
FT ZN_FING 8133 8155 C2H2-type 251.
FT ZN_FING 8163 8185 C2H2-type 252.
FT ZN_FING 8193 8215 C2H2-type 253.
FT ZN_FING 8223 8245 C2H2-type 254.
FT ZN_FING 8253 8275 C2H2-type 255.
FT ZN_FING 8283 8305 C2H2-type 256.
FT ZN_FING 8313 8335 C2H2-type 257.
FT ZN_FING 8343 8365 C2H2-type 258.
FT ZN_FING 8373 8395 C2H2-type 259.
FT ZN_FING 8403 8425 C2H2-type 260.
FT ZN_FING 8433 8455 C2H2-type 261.
FT ZN_FING 8463 8485 C2H2-type 262.
FT ZN_FING 8493 8515 C2H2-type 263.
FT ZN_FING 8523 8545 C2H2-type 264.
FT ZN_FING 8553 8575 C2H2-type 265.
FT ZN_FING 8583 8605 C2H2-type 266.
FT ZN_FING 8613 8635 C2H2-type 267.
FT ZN_FING 8643 8665 C2H2-type 268.
FT ZN_FING 8673 8695 C2H2-type 269.
FT ZN_FING 8703 8725 C2H2-type 270.
FT ZN_FING 8733 8755 C2H2-type 271.
FT ZN_FING 8763 8785 C2H2-type 272.
FT ZN_FING 8793 8815 C2H2-type 273.
FT ZN_FING 8823 8845 C2H2-type 274.
FT ZN_FING 8853 8875 C2H2-type 275.
FT ZN_FING 8883 8905 C2H2-type 276.
FT ZN_FING 8913 8935 C2H2-type 277.
FT ZN_FING 8943 8965 C2H2-type 278.
FT ZN_FING 8973 8995 C2H2-type 279.
FT ZN_FING 9003 9025 C2H2-type 280.
FT ZN_FING 9033 9055 C2H2-type 281.
FT ZN_FING 9063 9085 C2H2-type 282.
FT ZN_FING 9093 9115 C2H2-type 283.
FT ZN_FING 9123 9145 C2H2-type 284.
FT ZN_FING 9153 9175 C2H2-type 285.
FT ZN_FING 9183 9205 C2H2-type 286.
FT ZN_FING 9213 9235 C2H2-type 287.
FT ZN_FING 9243 9265 C2H2-type 288.
FT ZN_FING 9273 9295 C2H2-type 289.
FT ZN_FING 9303 9325 C2H2-type 290.
FT ZN_FING 9333 9355 C2H2-type 291.
FT ZN_FING 9363 9385 C2H2-type 292.
FT ZN_FING 9393 9415 C2H2-type 293.
FT ZN_FING 9423 9445 C2H2-type 294.
FT ZN_FING 9453 9475 C2H2-type 295.
FT ZN_FING 9483 9505 C2H2-type 296.
FT ZN_FING 9513 9535 C2H2-type 297.
FT ZN_FING 9543 9565 C2H2-type 298.
FT ZN_FING 9573 9595 C2H2-type 299.
FT ZN_FING 9603 9625 C2H2-type 300.
FT ZN_FING 9633 9655 C2H2-type 301.
FT ZN_FING 9663 9685 C2H2-type 302.
FT ZN_FING 9693 9715 C2H2-type 303.
FT ZN_FING 9723 9745 C2H2-type 304.
FT ZN_FING 9753 9775 C2H2-type 305.
FT ZN_FING 9783 9805 C2H2-type 306.
FT ZN_FING 9813 9835 C2H2-type 307.
FT ZN_FING 9843 9865 C2H2-type 308.
FT ZN_FING 9873 9895 C2H2-type 309.
FT ZN_FING 9903 9925 C2H2-type 310.
FT ZN_FING 9933 9955 C2H2-type 311.
FT ZN_FING 9963 9985 C2H2-type 312.
FT ZN_FING 9993 10015 C2H2-type 313.
FT ZN_FING 10023 10045 C2H2-type 314.
FT ZN_FING 10053 10075 C2H2-type 315.
FT ZN_FING 10083 10105 C2H2-type 316.
FT ZN_FING 10113 10135 C2H2-type 317.
FT ZN_FING 10143 10165 C2H2-type 318.
FT ZN_FING 10173 10195 C2H2-type 319.
FT ZN_FING 10203 10225 C2H2-type 320.
FT ZN_FING 10233 10255 C2H2-type 321.
FT ZN_FING 10263 10285 C2H2-type 322.
FT ZN_FING 10293 10315 C2H2-type 323.
FT ZN_FING 10323 10345 C2H2-type 324.
FT ZN_FING 10353 10375 C2H2-type 325.
FT ZN_FING 10383 10405 C2H2-type 326.
FT ZN_FING 10413 10435 C2H2-type 327.
FT ZN_FING 10443 10465 C2H2-type 328.
FT ZN_FING 10473 10495 C2H2-type 329.
FT ZN_FING 10503 10525 C2H2-type 330.
FT ZN_FING 10533 10555 C2H2-type 331.
FT ZN_FING 10563 10585 C2H2-type 332.
FT ZN_FING 10593 10615 C2H2-type 333.
FT ZN_FING 10623 10645 C2H2-type 334.
FT ZN_FING 10653 10675 C2H2-type 335.
FT ZN_FING 10683 10705 C2H2-type 336.
FT ZN_FING 10713 10735 C2H2-type 337.
FT ZN_FING 10743 10765 C2H2-type 338.
FT ZN_FING 10773 10795 C2H2-type 339.
FT ZN_FING 10803 10825 C2H2-type 340.
FT ZN_FING 10833 10855 C2H2-type 341.
FT ZN_FING 10863 10885 C2H2-type 342.
FT ZN_FING 10893 10915 C2H2-type 343.
FT ZN_FING 10923 10945 C2H2-type 344.
FT ZN_FING 10953 10975 C2H2-type 345.
FT ZN_FING 10983 11005 C2H2-type 346.
FT ZN_FING 11013 11035 C2H2-type 347.
FT ZN_FING 11043 11065 C2H2-type 348.
FT ZN_FING 11073 11095 C2H2-type 349.
FT ZN_FING 11103 11125 C2H2-type 350.
FT ZN_FING 11133 11155 C2H2-type 351.
FT ZN_FING 11163 11185 C2H2-type 352.
FT ZN_FING 11193 11215 C2H2-type 353.
FT ZN_FING 11223 11245 C2H2-type 354.
FT ZN_FING 11253 11275 C2H2-type 355.
FT ZN_FING 11283 11305 C2H2-type 356.
FT ZN_FING 11313 11335 C2H2-type 357.
FT ZN_FING 11343 11365 C2H2-type 358.
FT ZN_FING 11373 11395 C2H2-type 359.
FT ZN_FING 11403 11425 C2H2-type 360.
FT ZN_FING 11433 11455 C2H2-type 361.
FT ZN_FING 11463 11485 C2H2-type 362.
FT ZN_FING 11493 11515 C2H2-type 363.
FT ZN_FING 11523 11545 C2H2-type 364.
FT ZN_FING 11553 11575 C2H2-type 365.
FT ZN_FING 11583 11605 C2H2-type 366.
FT ZN_FING 11613 11635 C2H2-type 367.
FT ZN_FING 11643 11665 C2H2-type 368.
FT ZN_FING 11673 11695 C2H2-type 369.
FT ZN_FING 11703 11725 C2H2-type 370.
FT ZN_FING 11733 11755 C2H2-type 371.
FT ZN_FING 11763 11785 C2H2-type 372.
FT ZN_FING 11793 11815 C2H2-type 373.
FT ZN_FING 11823 11845 C2H2-type 374.
FT ZN_FING 11853 11875 C2H2-type 375.
FT ZN_FING 11883 11905 C2H2-type 376.
FT ZN_FING 11913 11935 C2H2-type 377.
FT ZN_FING 11943 11965 C2H2-type 378.
FT ZN_FING 11973 11995 C2H2-type 379.
FT ZN_FING 12003 12025 C2H2-type 380.
FT ZN_FING 12033 12055 C2H2-type 381.
FT ZN_FING 12063 12085 C2H2-type 382.
FT ZN_FING 12093 12115 C2H2-type 383.
FT ZN_FING 12123 12145 C2H2-type 384.
FT ZN_FING 12153 12175 C2H2-type 385.
FT ZN_FING 12183 12205 C2H2-type 386.
FT ZN_FING 12213 12235 C2H2-type 387.
FT ZN_FING 12243 12265 C2H2-type 388.
FT ZN_FING 12273 12295 C2H2-type 389.
FT ZN_FING 12303 12325 C2H2-type 390.
FT ZN_FING 12333 12355 C2H2-type 391.
FT ZN_FING 12363 12385 C2H2-type 392.
FT ZN_FING 12393 12415 C2H2-type 393.
FT ZN_FING 12423 12445 C2H2-type 394.
FT ZN_FING 12453 12475 C2H2-type 395.
FT ZN_FING 1
```

```

RA Miranda A., Mungall C.J., Munoz J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY18408; AAA48437.1; -.
DR FlyBase: FBgn035879; CG7112.
DR InterPro: IPR011036; PH related.
DR InterPro: IPR00195; RabGAP_TBC.
DR Pfam: PF00566; TBC; 1.
DR SMART: SM00164; TBC; 1.
DR PROSITE: PSS0086; TBC_RABGAP; 1.
SQ SEQUENCE 1005 AA; 113317 MW; 58C70A8326D2073A CRC64;

Query Match 43.5%; Score 47; DB 2; Length 1005;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 PTLRQWLEGRPRKN 19
Db 496 PILEWDESKPRKN 509

RESULT 35
ID Q9VSI2 PRELIMINARY; PRT; 1005 AA.
AC Q9VSI2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE CG7112-PA.
GN ORFNames=CG7112;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoeklin R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blaise J.R., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abell J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
RA Balcer R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegam C.,
RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Meckllov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palnert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shie B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodgerf, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

```

```

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoeklin R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.B., Myers B.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernier B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB003555; AAF50437.2; -.
DR FlyBase: FBgn035879; CG7112.
DR InterPro: IPR011036; PH related.
DR InterPro: IPR00195; RabGAP_TBC.
DR Pfam: PF00566; TBC; 1.
DR SMART: SM00164; TBC; 1.
DR PROSITE: PSS0086; TBC_RABGAP; 1.
SQ SEQUENCE 1005 AA; 113287 MW; 59DF5B4F840E2A55 CRC64;

Query Match 43.5%; Score 47; DB 2; Length 1005;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 PTLRQWLEGRPRKN 19
Db 496 PILEWDESKPRKN 509

RESULT 36
ID Q9P3E2 PRELIMINARY; PRT; 1171 AA.
AC Q9P3E2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Related to transport protein USO1.
GN Name=B13118.10;

```

```

OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algen V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390189; CAB99171.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008865; P:protein transporter activity; IEA.
DR GO; GO:0006886; P:intracellular protein transport; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR006955; USOL_P115_C.
DR InterPro; IPR006953; USOL_P115_head.
DR Pfam; PF04871; USOL_P115_C; 1.
DR Pfam; PF04869; USOL_P115_head; 1.
SQ SEQUENCE 1171 AA; 131632 MW; 33DF50E5931ED060 CRC64;

Query Match 43.5%; Score 47; DB 2; Length 1171;
Best Local Similarity 42.1%; Pred. No. 2.9e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 GNADGPTLRQWLREKRRKXN 19
DB 259 GSTDGEVQAQWAEORNRN 277

RESULT 37
Q9L1W0 PRELIMINARY; PRT; 522 AA.
ID Q9L1W0
AC Q9L1W0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to an Arabidopsis thaliana chromosome BAC genomic
DE sequence.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Heising Y.C., Chow T., Chen C., Wu H., Chu Y., Liu S.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001111; BAA90509.1; -.
DR Gramene; Q9L1W0; -.
SQ SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 43.1%; Score 46.5; DB 2; Length 522;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 5; Gaps 3;

QY 1 GNADG---PTLRQWL-EGRRP 17
DB 452 GVADGCIWPA-RQWLREGRRP 471

RESULT 38
EP42 HUMAN STANDARD; PRT; 690 AA.
ID EP42 HUMAN
AC P16452;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)

```

```

DE (P4.2).
GN Name=EP42; Synonym=E42P;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Reticulocytes;
RA MEDLINE=91271288; PubMed=2052563;
RX Korsgren C., Cohen C.M.;
RT "Organization of the gene for human erythrocyte membrane protein 4.2: structural similarities with the gene for the a subunit of factor XIII."
RL Proc. Natl. Acad. Sci. U.S.A. 88:4840-4844(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (ISOFORM SHORT).
RC TISSUE=Reticulocytes;
RA MEDLINE=90138679; PubMed=2300550;
RX Korsgren C., Lawler J., Lambert S., Speicher D., Cohen C.M.;
RT "Complete amino acid sequence and homologies of human erythrocyte membrane protein band 4.2."
RL Proc. Natl. Acad. Sci. U.S.A. 87:613-617(1990).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=Reticulocytes;
RA MEDLINE=90138995; PubMed=1689063;
RX Sung L.A., Chien S., Chang L.-S., Lambert K., Bliss S.A.,
RA Bouhasseira E.E., Nagel R.L., Schwartz R.S., Rydicki A.C.;
RT "Molecular cloning of human protein 4.2: a major component of the erythrocyte membrane."
RL Proc. Natl. Acad. Sci. U.S.A. 87:955-959(1990).
RN [4]
RP MYRISTOYLATION.
RX MEDLINE=92184834; PubMed=1544941;
RA Risinger M.A., Docimas R.M., Cohen C.M.;
RT "Human erythrocyte protein 4.2, a high copy number membrane protein, is N-myristylated."
RL J. Biol. Chem. 267:5680-5685(1992).
RN [5]
RP PHOSPHORYLATION SITE SER-247.
RX MEDLINE=93271204; PubMed=8499466; DOI=10.1016/0005-2736(93)90156-T;
RA Docimas E., Speicher D.W., Gupatroy B., Cohen C.M.;
RT "Structural domain mapping and phosphorylation of human erythrocyte pallidin (band 4.2)."
RL Biochim. Biophys. Acta 1148:19-29(1993).
RN [6]
RP VARIANT HS THR-111.
RX MEDLINE=92216098; PubMed=1558976;
RA Bouhasseira E.E., Schwartz R.S., Yawata Y., Ata K., Kanazaki A.,
RA Qiu J.J.-H., Nagel R.L., Rydicki A.C.;
RT "An alanine-to-threonine substitution in protein 4.2 cDNA is associated with a Japanese form of hereditary hemolytic anemia (protein 4.2 Nippon)."
RL Blood 79:1846-1854(1992).
RN [7]
RP VARIANT HS THR-111.
RX MEDLINE=95118828; PubMed=7819064;
RA Takaoka Y., Ideguchi H., Matsuda M., Sakamoto N., Takeuchi T.,
RA Fukunaki Y.;
RT "A novel mutation in the erythrocyte protein 4.2 gene of Japanese patients with hereditary spherocytosis (protein 4.2 Fukuoka)."
RL Br. J. Haematol. 88:527-533(1994).
RN [8]
RP VARIANT HS GLN-279.
RX MEDLINE=95290393; PubMed=7772513;
RA Hayette S., Morle L., Bozon M., Ghanem A., Risinger M., Korsgren C.,
RA Tanner M.J.A., Fattoum S., Cohen C.M., Delaunay J.;
RT "A point mutation in the protein 4.2 gene (allele 4.2 Tozeur) associated with hereditary haemolytic anaemia."
RL Br. J. Haematol. 89:762-770(1995).
CC -!- FUNCTION: Probably plays an important role in the regulation of erythrocyte shape and mechanical properties.

```

CC -1- SUBUNIT: Oligomer. Interacts with the cytoplasmic domain of  
 CC SLC4A1/band 3 anion transport protein.  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (cytoplasmic surface of  
 CC erythrocyte membranes) and cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name-Short;  
 CC IsoId=P16452-1; Sequence=Displayed;  
 CC Note=Major isoform;  
 CC Name=Long;  
 CC IsoId=P16452-2; Sequence=VSP\_006416;  
 CC -1- PFM: Both cAMP-dependent kinase (CAK) and another kinase present  
 CC in the red blood cells seem to be able to phosphorylate EPR42.  
 CC -1- DISEASE: Defects in EPR42 are a cause of hereditary spherocytosis  
 CC (HS) [MIM:177070], a hematologic disorder leading to chronic  
 CC hemolytic anemia and characterized by numerous abnormally shaped  
 CC erythrocytes which are generally spheroidal. Absence of band 4.2  
 CC associated with spur or target erythrocytes has also been  
 CC reported.  
 CC -1- MISCELLANEOUS: The substitution of an Ala for a Cys in the active  
 CC site may be responsible for the lack of transglutaminase activity  
 CC of band 4.2.  
 CC -1- SIMILARITY: Belongs to the transglutaminase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; M60298; AAA74589.1; -;  
 DR EMBL; L06519; AAA52385.1; -;  
 DR EMBL; L06447; AAA52385.1; JOINED.  
 DR EMBL; L06448; AAA52385.1; JOINED.  
 DR EMBL; L06449; AAA52385.1; JOINED.  
 DR EMBL; L06450; AAA52385.1; JOINED.  
 DR EMBL; L06512; AAA52385.1; JOINED.  
 DR EMBL; L06511; AAA52385.1; JOINED.  
 DR EMBL; L06513; AAA52385.1; JOINED.  
 DR EMBL; L06515; AAA52385.1; JOINED.  
 DR EMBL; L06516; AAA52385.1; JOINED.  
 DR EMBL; L06517; AAA52385.1; JOINED.  
 DR EMBL; L06518; AAA52385.1; JOINED.  
 DR EMBL; M29399; AAA5798.1; -;  
 DR EMBL; M30646; AAA56402.1; -;  
 DR EMBL; M30647; AAA56401.1; -;  
 DR PIR; A39707; A39707.  
 DR HSSP; P52181; 1G0D.  
 DR Genew; HGNC:3381; EPR42.  
 DR MIM; 177070; -;  
 DR GO; GO:0005856; C:cytoskeleton; TAS.  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0005524; F:ATP binding; TAS.  
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.  
 DR InterPro; IPR001102; Glutaminase.  
 DR InterPro; IPR008958; Transglut\_C.  
 DR InterPro; IPR002931; Transglutase\_like.  
 DR Pfam; PF00927; Transglut\_C\_2.  
 DR Pfam; PF01841; Transglut\_core; 1.  
 DR Pfam; PF01841; Transglut\_N; 1.  
 DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.  
 KW Alternative splicing; Cell shape; Cytoskeleton;  
 KW Direct protein sequencing; Disease mutation; Erythrocyte maturation;  
 KW Hereditary hemolytic anemia; Lipoprotein; Myristate; Phosphorylation;  
 KW Structural protein.  
 KW INIT\_MET 0  
 FT SITE 30 38 By similarity. (By similarity).  
 FT LIPID 1 1 N-myristoyl glycine.  
 FT MOD\_RES 247 247 Phosphoserine (by PKA) (Probable).  
 FT VARSPLIC 2 2 Q->QGRPSQSTSLAGLVAPAPASPVFKSGMD (in  
 FT isoform Long).

FT VARIANT 111 111 /FTId=VSP\_006416.  
 FT FT A->T (in HS; Nippon/Rukoka).  
 FT VARIANT 279 279 /FTId=VAR\_007482.  
 FT FT R->Q (in HS; Tozeur).  
 FT CONFLICT 334 339 /FTId=VAR\_012268.  
 FT CONFLICT 349 349 TRPALP->KRGIPC (in Ref. 3).  
 FT CONFLICT 375 375 D->H (in Ref. 3).  
 FT CONFLICT 375 375 V->L (in Ref. 3).  
 SQ SEQUENCE 690 AA; 76841 MW; C6E605E59A0A7A8B CRC64;  
 Query Match 43.1%; Score 46.5; DB 1; Length 690;  
 Best Local Similarity 76.9%; Pred. No. 1.9e+02;  
 Matches 10; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 Qy 6 PTLRQMLEGR-RP 17  
 Db 249 PILRQWLTRGRGP 261  
 PTLRQMLEGR-RP 17  
 PILRQWLTRGRGP 261  
 RESULT 39  
 Q95QV6 PRELIMINARY; PRT; 103 AA.  
 ID Q95QV6;  
 AC Q95QV6;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Hypothetical protein C18A3.5.  
 GN Name=C18A3.5; ORFNames=C18A3.5;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Peleoderinae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG Wormbase Consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Hallsworth K.;  
 RT "The sequence of C. elegans cosmid C18A3.";  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG Wormbase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U28944; AKK68193.1; -;  
 DR Wormbase; WBGene0015943; C18A3.5.  
 DR WormPep; C18A3.5c; CE27710.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; RRM\_1; 1.  
 DR PROSITE; PS50102; RRM; 1.  
 KW Hypothetical protein.



SQ SEQUENCE 103 AA; 11420 MW; 6D3A1877857E5E64 CRC64;

Query Match 42.6%; Score 46; DB 2; Length 103;

Best Local Similarity 72.7%; Pred. No. 30;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNADGPTLRQW 11  
Db 92 GNQSTPTLRQW 102

RESULT 40

Q9NMW3 PRELIMINARY; PRT; 157 AA.

AC Q9NMW3; (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, last sequence update)  
DE 01-OCT-2002 (T-EMBLrel. 22, last annotation update)  
DE Hypothetical protein FLJ10043.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole embryo;  
RX PubMed=14702039; DOI=10.1038/ng1285;  
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,  
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
RA Sudo H., Hoshiro T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
RA Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M., Yamazaki M.,  
RA Ninomiya K., Iishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
RA Tanai H., Kimura M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,  
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hoshino T., Nomura Y.,  
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
RA Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
RA Musahiro K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Taniguchi A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno H., Yamashita R.,  
RA Nakai K., Tada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RT RT CDNAS."  
RT "Complete sequencing and characterization of 21,243 full-length human  
DR ENBL: AK00905; BAA91418.1;  
SQ SEQUENCE 157 AA; 17352 MW; 2B1C9C747758C2D23 CRC64;

Query Match 42.6%; Score 46; DB 2; Length 157;

Best Local Similarity 61.5%; Pred. No. 47;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 PTLRWLEGRPK 18  
Db 15 PTLRWLEGRPK 27

RESULT 41

Q9RKRP9 PRELIMINARY; PRT; 243 AA.

AC Q9RKRP9; (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, last annotation update)  
DE Hypothetical protein SC02279.  
GN ORFNames=SC02279.25;  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
OC Streptomycetaceae; Streptomyces.  
NC NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=21996410; PubMed=1200953; DOI=10.1038/417141a;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
RA Rabbinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2)";  
RL Nature 417:141-147(2002).  
DR EMBL; AL939112; CAB61725.1; -.  
DR PIR; T50588; T50588.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 243 AA; 26559 MW; 13584D7A81A0E990 CRC64;

Query Match 42.6%; Score 46; DB 2; Length 243;  
Best Local Similarity 72.7%; Pred. No. 75;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 LKOWLEGRPK 18  
Db 15 LKOWLEGRPK 25

RESULT 42

Q9W2V0 PRELIMINARY; PRT; 249 AA.

AC Q9W2V0; (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, last annotation update)  
DE Hypothetical protein OSJNB0076H04.22 (Putative reverse  
DE transcriptase).  
GN Name=OSJNB0076H04.22; ORFNames=OSJNB0022D10.9;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriarthridae; Oryzaceae; Oryza.  
NC NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffet K.S., Hill J.N.,  
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,  
RA Teitler T., Riggs F., Hsiao J., Ziemann V., Blunt S., Pai G.,  
RA Vanaken S.E., Utterback T.R., Feldlyum T.V., Kalb E., Quackenbush J.,  
RA Salzberg S.L., White O., Fraser C.M.,  
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Buell R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA The Rice Chromosome 10 Sequencing Consortium;  
RT "In-depth view of structure, activity, and evolution of rice  
RT chromosome 10";  
RL Science 300:1566-1569(2003).



```

RN (4)
RP SEQUENCE FROM N.A.
RA Buehl C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC093093; AA58152.1; -
DR EMBL; AB017082; AAP53273.1; -
DR Gramene; O7XFE3; -
DR GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.
KM Hypothetical protein; RNA-directed DNA polymerase.
SQ SEQUENCE 249 AA; 28243 MW; 05E3D2406DFA7C7C CRC64;

Query Match 42.6%; Score 46; DB 2; Length 249;
Best Local Similarity 45.0%; Pred. No. 77;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Db 120 DGNTRFWDMSAMIDGRRPKD 139

RESULT 43
07G731 PRELIMINARY; PRT; 249 AA.
ID 07G731;
AC 07G731;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative reverse transcriptase.
GN Name=OSUNBA002D10.9;
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
OC NCB1_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Saeki C., Henry D., Oates R., Simmons J.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC093402; AAL79345.1; -
DR GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.
KM RNA-directed DNA polymerase.
SQ SEQUENCE 249 AA; 28243 MW; 05E3D2406DFA7C7C CRC64;

Query Match 42.6%; Score 46; DB 2; Length 249;
Best Local Similarity 45.0%; Pred. No. 77;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Cy 4 DGPTLR---QWLSGRPRKN 19
Db 120 DGNTRFWDMSAMIDGRRPKD 139

RESULT 44
09S1Q2 PRELIMINARY; PRT; 357 AA.
ID 09S1Q2;
AC 09S1Q2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein SCO0239.
GN ORFNames=SCJ39A.18c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OK NCB1_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=A3(12) / M145;
RC MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bertley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,

```

```

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajadream M.A., Rutherford K.M., Ruter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wenzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)." (2002).
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB53279.1; -
DR PIR; T37154; T37154.
DR InterPro; IPR011009; Kinase like.
DR Complete proteome; Hypothetical protein.
KM SEQUENCE 357 AA; 39139 MW; 731696F25D03B4AF CRC64;

Query Match 42.6%; Score 46; DB 2; Length 357;
Best Local Similarity 41.2%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Cy 1 GNADPTLRQWLSGRPR 17
Db 290 GTERGAFLREWDGHOP 306

RESULT 45
082Q03 PRELIMINARY; PRT; 362 AA.
ID 082Q03;
AC 082Q03;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Putative regulatory protein.
GN OrderedLocustNames=SAV472;
OS Streptomyces avermitilis.
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OK NCB1_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680;
RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680;
RC MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005023; BAC68182.1; -
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000005; HTHArac.
DR Pfam; PF00165; HTH_Arac; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KM Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 362 AA; 39245 MW; 4D3162C6E6A56CA0 CRC64;

Query Match 42.6%; Score 46; DB 2; Length 362;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

Fri Sep 2 09:00:10 2005

Qy	3	ADGPTLRQWLEGRPK	18
	:	:	:
Db	117	AECTVNGWIRGRRLK	132

Search completed: September 1, 2005, 16:21:01  
Job time : 75.6691 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 87.3453 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-8  
Perfect score: 114  
Sequence: 1 GGCADGPTLRKRWISFCGK 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues  
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_GeneSeq\_16Dec04:\*  
1: geneSeqp19808:\*  
2: geneSeqp19908:\*  
3: geneSeqp20008:\*  
4: geneSeqp20018:\*  
5: geneSeqp20028:\*  
6: geneSeqp20038:\*  
7: geneSeqp20048:\*  
8: geneSeqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	19	2	AAW09458 Thrombopo
2	114	100.0	19	2	AAW33025 Thrombopo
3	114	100.0	19	4	AAU25822 Human thr
4	109	95.6	18	2	AAW09456 Thrombopo
5	109	95.6	18	2	AAW33023 Thrombopo
6	109	95.6	18	3	AAU25820 Thrombopo
7	109	95.6	18	4	AAU25820 Human thr
8	109	95.6	18	5	ABU72906 TPO mimet
9	109	95.6	18	5	ADJ73058 TPO mimet
10	109	95.6	18	8	ADJ52693 CHI delet
11	109	95.6	18	8	ADJ51654 CHI delet
12	85	74.6	14	2	AAW09466 Thrombopo
13	85	74.6	14	2	AAW09462 Thrombopo
14	85	74.6	14	2	AAW09465 Thrombopo
15	85	74.6	14	2	AAW09482 Thrombopo
16	85	74.6	14	2	AAW33031 Thrombopo
17	85	74.6	14	2	AAW36633 Thrombopo
18	85	74.6	14	2	AAW33029 Thrombopo
19	85	74.6	14	2	AAW35401 Thrombopo
20	85	74.6	14	2	AAW36647 Thrombopo
21	85	74.6	14	2	AAW35400 Thrombopo
22	85	74.6	14	2	AAW33032 Thrombopo
23	85	74.6	14	3	AAU25826 Human thr
24	85	74.6	14	4	AAU25826 Human thr
25	85	74.6	14	4	AAU25852 Human thr

26	85	74.6	14	4	AAU25866
27	85	74.6	14	5	ABU72900
28	85	74.6	14	7	ADJ73051
29	85	74.6	14	8	ADJ52686
30	85	74.6	14	8	ADJ51647
31	76	66.7	13	2	AAW09467
32	76	66.7	13	2	AAW35399
33	76	66.7	13	2	AAW35417
34	76	66.7	13	2	AAW33033
35	76	66.7	13	2	AAW35413
36	76	66.7	13	2	AAW35406
37	76	66.7	13	2	AAW35422
38	76	66.7	13	2	AAW35397
39	76	66.7	13	4	AAU25997
40	76	66.7	14	4	AAU25984
41	76	66.7	14	2	AAW35398
42	76	66.7	14	2	AAW35396
43	76	66.7	14	2	AAW35402
44	76	66.7	14	4	AAU25987
45	76	66.7	14	4	AAU25983
46	76	66.7	14	4	AAU25985
47	72	63.2	12	2	AAW35423
48	72	63.2	12	4	AAU26000
49	67	58.8	13	2	AAW35404
50	67	58.8	13	2	AAW35405
51	67	58.8	13	4	AAU25994
52	67	58.8	13	4	AAU25991
53	67	58.8	13	4	AAU25990
54	67	58.8	14	2	AAW35412
55	67	58.8	14	2	AAW35407
56	67	58.8	14	2	AAW35408
57	67	58.8	14	2	AAW35403
58	67	58.8	14	4	AAU25993
59	67	58.8	14	4	AAU25989
60	67	58.8	14	4	AAU25995
61	67	58.8	14	4	AAU25992
62	67	58.8	14	4	AAU25986
63	67	58.8	14	4	AAU25988
64	67	58.8	25	4	AAU26042
65	67	58.8	25	8	ADU72531
66	66	57.9	11	2	AAW35425
67	66	57.9	11	4	AAU26001
68	66	57.9	25	7	ADN59740
69	65	57.0	13	4	AAU26041
70	64	56.1	14	3	AAU26017
71	64	56.1	14	5	ABU72903
72	64	56.1	14	8	ADJ52689
73	64	56.1	10	2	AAW35427
74	60	52.6	10	4	AAU26002
75	60	52.6	18	7	ADN59680
76	60	52.6	22	7	ADN59839
77	59	51.8	25	7	ADN59744
78	59	51.8	12	8	ADU72530
79	57	50.0	13	4	AAU26039
80	57	50.0	13	8	ADU72529
81	57	50.0	13	8	ADU72528
82	57	50.0	14	2	AAW6732
83	57	50.0	14	4	AAU26040
84	57	50.0	16	2	AAW09464
85	57	50.0	16	2	AAW33329
86	57	50.0	16	3	AAU26019
87	57	50.0	16	3	AAU25829
88	57	50.0	16	4	AAU25828
89	57	50.0	16	5	ABU72905
90	57	50.0	16	7	ADU73057
91	57	50.0	16	8	ADJ52692
92	57	50.0	16	8	ADJ51653
93	56.5	49.6	23	7	ADN59778
94	56.5	49.6	41	7	ADN59816
95	56.5	49.6	41	7	ADN59772
96	56.5	49.6	46	7	ADN59790
97	56.5	49.6	46	7	ADN59784
98	56	49.1	13	3	AAU26015

AAU25866	Human thr
ABU72900	TPO mimet
ADJ73051	TPO mimet
ADJ52686	CHI delet
ADJ51647	CHI delet
AAW09467	Thrombopo
AAW35399	Thrombopo
AAW35417	Thrombopo
AAW33033	Thrombopo
AAW35413	Thrombopo
AAW35406	Thrombopo
AAW35422	Thrombopo
AAW35397	Thrombopo
AAU25997	Human thr
AAU25984	Human thr
AAW35398	Thrombopo
AAW35396	Thrombopo
AAW35402	Thrombopo
AAU25987	Human thr
AAU25983	Human thr
AAU25985	Human thr
AAW35423	Thrombopo
AAU26000	Human thr
AAW35404	Thrombopo
AAW35405	Thrombopo
AAU25994	Human thr
AAU25991	Human thr
AAU25990	Human thr
AAW35412	Thrombopo
AAW35407	Thrombopo
AAW35408	Thrombopo
AAW35403	Thrombopo
AAU25993	Human thr
AAU25989	Human thr
AAU25995	Human thr
AAU25992	Human thr
AAU25986	Human thr
AAU25988	Human thr
AAU26042	Human thr
ADU72531	TPO mimet
AAW35425	Thrombopo
AAU26001	Human thr
ADN59740	Thrombopo
AAU26041	Human thr
AAU26017	TPO-mimet
ABU72903	TPO mimet
ADJ52689	CHI delet
ADJ51650	CHI delet
AAW35427	Thrombopo
AAU26002	Human thr
ADN59680	Thrombopo
ADN59839	TMP Pepti
ADN59744	Thrombopo
ADU72530	TPO mimet
AAU26039	Human thr
ADU72529	TPO mimet
ADU72528	TPO mimet
AAW6732	Peptide c
AAU26040	Human thr
AAW09464	Thrombopo
AAW33329	Thrombopo
AAU26019	TPO-mimet
AAU25829	Human thr
AAU25828	Human thr
ABU72905	TPO mimet
ADJ73057	TPO mimet
ADJ52692	CHI delet
ADJ51653	CHI delet
ADN59778	Peptide-v
ADN59816	Peptide-v
ADN59772	Peptide-v
ADN59790	Peptide-v
ADN59784	Peptide-v
AAU26015	TPO-mimet

99 56 49.1 13 5 ABB72901  
100 56 49.1 13 7 ADJ73054

Abb72901 TPO mimet  
Adj73054 TPO mimet

## ALIGNMENTS

## RESULT 1

AAW09458

ID AAW09458 standard; protein; 19 AA.

AAW09458;

10-SEP-1997 (first entry)

Thrombopoietin receptor binding compound peptide.

Haematology; thrombocytopenia; TPO; TR; proliferation;

bone marrow transfusion; chemotherapy; radiation therapy.

Synthetic.

Location/Qualifiers

Key Misc-difference 1..19

/note="Preferably linkages are selected from: -

CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -C(O)NR6

; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is

lower alkyl"]

Modified-site

1 /note="Preferably N-terminus is selected from: -NRR1; -

NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;

benzyloxycarbonyl-NH; benzyloxycarbonyl-NH with 1-3

substitutions on the phenyl ring selected from lower

alkyl, lower alkoxy, chloro, bromo; where R and R1 are

independently selected from hydrogen and lower alkyl"]

Modified-site

19 /note="Preferably C-terminus is -C(O)R2 where R2 is

selected from hydroxy, lower alkoxy, and -NRR4, where R3

and R4 are independently selected from hydrogen and lower

alkyl, and where the nitrogen atom of the -NRR4 group

can optionally be the amine group of the N-terminus of

the peptide forming a cyclic peptide"

WO9640189-A1.

19-DEC-1996.

05-JUN-1996; 96WO-US008998.

07-JUN-1995; 95US-00472371.

07-JUN-1995; 95US-00473604.

07-JUN-1995; 95US-00476168.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00484090.

07-JUN-1995; 95US-00485301.

(GLAX ) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-051883/05.

Thrombopoietin receptor-binding/activating peptide(s) and peptide

mimetic(s) - useful in treatment of haematological disorders, esp.

thrombocytopenia resulting from chemotherapy, etc.

Claim 18; Page 89; 106pp; English.

The present sequence is a compound which binds to thrombopoietin (TPO)

affinity to TR as expressed by an IC50 of no more than about 100 nmM. The

compound (especially if modified, see features table) can be used for

CC treating patients suffering from haematological disorders and  
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
CC marrow transfusions. The peptide may also be used to maintain the  
CC proliferation and growth of TPO-dependent cell lines and for use in  
CC biological research, for detecting TPO receptors on living cells

SQ Sequence 19 AA;

Query Match 100.0%; Score 114; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFGGK 19

Db 1 GGCADGPTLRWISFGGK 19

## RESULT 2

AAW33025

ID AAW33025 standard; peptide; 19 AA.

AAW33025;

11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;

haematological disorder; thrombocytopenia; diagnosis;

radiation therapy; bone marrow transfusion; cell culture.

signal transduction; receptor activation; cell culture.

Synthetic.

WO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAX ) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the

thrombopoietin receptor - useful in treatment of haematological

disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Claim 19; Page 89; 106pp; English.

The present peptide binds the thrombopoietin receptor (TR), has a

molecular weight of less than 8000 Da and a TR binding affinity as

expressed by an IC50 of no more than about 100 microm. It can be used to

treat disorders which are susceptible to treatment with a thrombopoietin

agonist, preferably haematological disorders and thrombocytopenia

resulting from chemotherapy, radiation therapy or bone marrow

transfusions. It can also be used diagnostically, e.g. to investigate the

mechanism of thrombopoietin signal transduction and receptor activation,

or to maintain the proliferation and growth of thrombopoietin dependent

cell lines

SQ Sequence 19 AA;

Query Match 100.0%; Score 114; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGCADGPTLRWISFCGK 19  
 |||||  
 Db 1 GGCADGPTLRWISFCGK 19

## RESULT 3

AAU25822  
 ID AAU25822 standard; peptide; 19 AA.

AAU25822;  
 DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #8.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iacti gene.

OS Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-05009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Podduturi S;  
 PI Yin Q;

XX WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 67-68; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 19 AA;

Query Match 100.0%; Score 114; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGCADGPTLRWISFCGK 19  
 |||||  
 Db 1 GGCADGPTLRWISFCGK 19

## RESULT 4

AAW09456  
 ID AAW09456 standard; protein; 18 AA.

AAW09456;  
 DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;  
 KW bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX Key

XX Misc-difference

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

Location/Qualifiers  
 1. 18  
 /note= "Preferably linkages are selected from: -CH2C(O)NR6  
 ; -NHC(O)NR; where R is hydrogen or lower alkyl and R6 is  
 lower alkyl"  
 18  
 /note= "Preferably N-terminus is selected from: -NRR1; -  
 NRC(O)R; -NRC(O)R; -NRC(O)R; -NRC(O)R; -NRC(O)R; -NRC(O)R;  
 benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3  
 substituents on the phenyl ring selected from lower  
 alkyl, lower alkoxy, chloro, bromo; where R and R1 are  
 independently selected from hydrogen and lower alkyl"  
 18  
 /note= "Preferably C-terminus is -C(O)R2 where R2 is  
 selected from hydroxy, lower alkoxy, and -NR3R4, where R3  
 and R4 are independently selected from hydrogen and lower  
 alkyl, and where the nitrogen atom of the -NR3R4 group  
 can optionally be the amine group of the N-terminus of  
 the peptide forming a cyclic peptide"

XX W09640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

XX 07-JUN-1995; 95US-00473604.

XX 07-JUN-1995; 95US-00476168.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00484090.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide  
 PT mimetic(s) - useful in treatment of haematological disorders, esp.  
 PT thrombocytopenia resulting from chemotherapy, etc.

XX Claim 18; Page 89; 106pp; English.

The present sequence is a compound which binds to thrombopoietin (TPO)  
 receptor (TR). It has a molecular weight of < 8000 Da, and a binding  
 affinity to TR as expressed by an IC50 of no more than about 100 nm. The  
 compound (especially if modified, see features table) can be used for  
 treating patients suffering from haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. The peptide may also be used to maintain the  
 CC proliferation and growth of TPO-dependent cell lines and for use in  
 CC biological research, for detecting TPO receptors on living cells  
 XX

Sequence 18 AA;

Query Match 95.6%; Score 109; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 Db 1 GGCADGPTLRWISFCGG 18

RESULT 5  
 ID AAW33023 standard; peptide; 18 AA.

XX AAW33023;

DT 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

KM signal transduction; receptor activation; cell culture.

OS Synthetic.

XX MO640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a  
 CC molecular weight of less than 8000 Da and a TR binding affinity as  
 CC expressed by an IC50 of no more than about 100 microm. It can be used to  
 CC treat disorders which are susceptible to treatment with a thrombopoietin  
 CC agonist, preferably haematological disorders and thrombocytopenia  
 CC resulting from chemotherapy, radiation therapy or bone marrow  
 CC transfusions. It can also be used diagnostically, e.g. to investigate the  
 CC mechanism of thrombopoietin signal transduction and receptor activation,  
 CC or to maintain the proliferation and growth of thrombopoietin dependent  
 CC cell lines

XX Sequence 18 AA;

Query Match 95.6%; Score 109; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18

Db 1 GGCADGPTLRWISFCGG 18

RESULT 6  
 ID AAB17020 standard; peptide; 18 AA.

XX AAB17020;

DT 31-OCT-2000 (first entry)

XX TPO-mimetic peptide sequence SEQ ID NO:76.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KM autoimmune disease; cytostatic; antiaesthetic; thrombolytic; VEGF;  
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KM inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KM vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KM thrombolysis; pharmaceutical.

OS Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGEN) AMGEN INC.

XX Feige U, Liu C, Cheatham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 220; 608pp; English.

XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)-d-P2, -(L1)-c-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antiaesthetic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombolysis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions,  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention

XX Sequence 18 AA;

Query Match 95.6%; Score 109; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 Db 1 GGCADGPTLRWISFCGG 18

RESULT 7  
AAU25820 standard; peptide; 18 AA.  
XX ID AAU25820 standard; peptide; 18 AA.  
XX AC AAU25820;  
XX DT 17-DEC-2001 (first entry)  
XX DE Human thrombopoietin receptor (TPO-R) activator peptide #6.  
XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
KW bone marrow transplantation; haematological disorder; platelet disorder;  
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lact gene.  
XX OS Homo sapiens.  
XX PN US6251864-B1.  
XX PD 26-JUN-2001.  
XX PF 01-MAR-2000; 2000US-00516704.  
XX PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
PR 07-JUN-1996; 96WO-US0009623.  
PR 15-AUG-1996; 96US-00699027.  
XX PA (GLAXO) GLAXO GROUP LTD.  
XX PI Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ,  
PI Balasubramanian P, Magstrom CR, Hendren RM, Deprince RB, Podduturi S;  
PI Yin Q;  
XX DR WPI; 2001-564142/63.  
XX PT Activating thrombopoietin receptors in cells, used to treat  
PT thrombocytopenia and hematological disorders, comprises contacting cells  
PT with peptides and peptide mimetics attached to hydrophilic polymers.  
XX PS Disclosure; Col 65-66; 128pp; English.  
XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
CC of activating thrombopoietin receptors in cells comprise contacting the  
CC cells with effective amounts of peptides and peptide mimetics attached to  
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
CC as that due to chemotherapy, radiation therapy or bone-marrow  
CC transplantation and to prevent thrombocytopenia in patients at risk. The  
CC sequences are used to treat and prevent haematological disorders  
CC including thrombocytopenia and platelet disorders. They are used in vitro  
CC as unique tools for understanding the biological role of thrombopoietin  
CC (TPO) and to develop other compounds that bind to and activate the TPO  
CC receptor. The peptides can be used to detect TPO receptors on living  
CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
CC in purified or natural biological materials. They may also be used for in  
CC situ staining, fluorescence-activated cell sorting, Western blotting and  
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
CC be used for in vitro expansion of megakaryocytes and their committed  
CC progenitors alone or in conjunction with additional cytokines  
XX SO Sequence 18 AA;  
Query Match 95.6%; Score 109; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGCGADGPTLRMISFCGG 18  
1 GGCGADGPTLRMISFCGG 18

RESULT 8  
ABB72906 standard; peptide; 18 AA.  
XX ID ABB72906 standard; peptide; 18 AA.  
XX AC ABB72906;  
XX DT 05-APR-2002 (first entry)  
XX DE TPO mimetic peptide SEQ ID NO:76.  
XX KW Modified peptide; mimetic; FC domain; fusion; immunoglobulin G; IgG; EPO;  
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
KW TNF-alpha inhibitor; Interleukin 1 antagonist; IL-1 antagonist; TNF;  
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
KW antianaemic; anorectic; antifertility; haemostatic; dermatological;  
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
KW sleep disorder; neurological degenerative disease; anaemia;  
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
KW Fanconi's syndrome.  
XX OS Homo sapiens.  
XX PN Synthetic.  
XX PD WO200183525-A2.  
XX PF 08-NOV-2001.  
XX PR 02-MAY-2001; 2001WO-US014310.  
XX PR 03-MAY-2000; 2000US-00563286.  
XX PA (AMGEN-) AMGEN INC.  
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;  
XX DR WPI; 2002-130313/17.  
XX PT Novel vehicle-peptide molecule or its multimers useful for treating  
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
PT diabetic retinopathy, obesity, sleep disorders and infertility.  
XX PS Claim 39; Page 44; 176pp; English.  
XX CC The present invention describes a vehicle-peptide molecule (I) or its  
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
CC cyostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
CC antianaemic, anorectic, antifertility, haemostatic, dermatological and  
CC neuroprotective activities. (I) can be used as a therapeutic or  
CC prophylactic agent as well as for screening purposes. (I) is useful for  
CC diagnosing diseases characterised by dysfunction of their associated  
CC protein of interest, for identifying normal or abnormal proteins of  
CC interest, as a part of diagnostic kit to detect the presence of their  
CC proteins of interest in a biological sample. Additionally, (I) is useful  
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
CC mimetic compounds are useful for treating disorders characterised by low  
CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
CC compounds are useful for treating conditions that involve an existing  
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35655 to ABB35777  
CC represent amino acid and nucleic acid sequences used in the  
CC exemplification of the present invention  
XX SO Sequence 18 AA;

Query Match 95.6%; Score 109; DB 5; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 |||||  
 DB 1 GGCADGPTLRWISFCGG 18

## RESULT 9

ID ADJ73058 standard; peptide; 18 AA.

XX ADJ73058;

XX 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SegID 512.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;

KM cardiovascular; infectious; malignant; neurologic disease; anaemia;

KM immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

XX TPO.

OS Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ ) CENTOCOR INC.

PI Heaven GA, Knight DM, Scallion BJ, Chrayeb J;

XX WPI; 2003-804237/75.

PT New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 512; 97pp; English.

CC This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is a TPO mimetic peptide sequence used to make a

CC mimetibody of the invention.

XX Sequence 18 AA;

Query Match 95.6%; Score 109; DB 7; Length 18;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 |||||  
 DB 1 GGCADGPTLRWISFCGG 18

RESULT 10  
 ADJ52693  
 ID ADJ52693 standard; peptide; 18 AA.

XX ADJ52693;

XX 06-MAY-2004 (first entry)

XX CHI deleted mimetibody-related peptide SegID512.

XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KM hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KM fungicide; gene therapy; immune disorder; cardiovascular disease;

KM arrhythmia; hypertension; heart failure; neurodegenerative;

KM multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KM cancerous condition; infectious disease; bacterial infection;

XX viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

XX WO2004002417-A2.

XX 08-JUN-2004.

XX 27-JUN-2003; 2003WO-US020347.

XX 28-JUN-2002; 2002US-0392431P.

XX (CENZ ) CENTOCOR INC.

PI Heaven GA, Knight DM, Chrayeb J, Scallion BJ, Neespor TC;

XX Kutowski KA;

XX WPI; 2004-082870/08.

PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX Claim 2; SEQ ID NO 512; 129pp; English.

CC This invention relates to CHI deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an immunosuppressive,

CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,

CC antibacterial, virucide or fungicide activity. In addition, the disclosed

CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody

CC is useful for diagnosing or treating a disease condition in a cell,

CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart

CC failure), or neurodegenerative (for example multiple sclerosis, dementia

CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or

CC fungal infection). The present sequence is that of a peptide which may be

CC used during the creation of a mimetibody of the invention.

XX Sequence 18 AA;

Query Match 95.6%; Score 109; DB 8; Length 18;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCADGPTLRWISFCGG 18  
 |||||  
 DB 1 GGCADGPTLRWISFCGG 18

## RESULT 11



ADJ51654  
ID ADJ51654 standard; peptide, 18 AA.  
XX  
AC ADJ51654;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE CH1 deleted mimetibody-related peptide SeqID512.  
XX  
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
KW antiallergic; muscular-Gen; cytostatic; antinflammatory; neuroleptic;  
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
KW obstructive disorder; haematologic disorder; immunological disorder;  
KW allergic disorder; infectious disorder; musculoskeletal disorder;  
KW oncological disorder; neurological disorder; nutritional disorder;  
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
KW renal disorder; pulmonary disorder.  
XX  
XX Unidentified.  
OS Synthetic.  
OS  
XX WO2004002424-A2.  
PN  
XX 08-JAN-2004.  
PD  
XX 30-JUN-2003; 2003WO-US020495.  
PF  
XX 28-JUN-2002; 2002US-0392431P.  
PR  
XX 19-SEP-2002; 2002US-0412144P.  
PR  
XX (CENZ ) CENTOCOR INC.  
PA  
XX Heaven GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;  
PI Kucloski KA;  
PI  
XX WPI; 2004-082872/08.  
DR  
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for  
PT diagnosing, preventing or treating cardiovascular, dermatologic,  
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
PT nutritional disorders.  
XX  
PS Claim 15, SEQ ID NO 512, 123pp; English.  
PS  
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an osteopathic,  
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
CC antinflammatory, neuroleptic, ophthalmological, nephrotropic or  
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
CC modulator or cytokine-agonist. The methods and compositions of the  
CC present invention are useful for the diagnosis, prevention and/or  
CC treatment of diseases or conditions associated with aberrant expression  
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,  
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
CC obstructive, haematologic, immunological, allergic, infectious,  
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
CC pediatric, psychiatric, renal or pulmonary disorders. The present  
CC sequence is that of a peptide which may be used during the creation of a  
XX mimetibody of the invention.  
XX  
XX Sequence 18 AA;  
Query Match 95.6%; Score 109; DB 8; Length 18;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GGCGDPTLRWISFCGG 18  
DB 1 GGCGDPTLRWISFCGG 18  
RESULT 12  
AAW09466  
ID AAW09466 standard; protein, 14 AA.  
XX  
AC AAW09466;  
XX  
DT 10-SEP-1997 (first entry)  
XX  
DE Thrombopoietin receptor binding compound cyclic peptide.  
KW Thrombopoietin receptor binding compound cyclic peptide.  
KW Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;  
KW bone marrow transfusion; chemotherapy; radiation therapy.  
XX  
XX Synthetic.  
OS  
XX Key Location/Qualifiers  
FH Disulfide-bond 1. 14  
FT Modified-site 1 /note= "In acetyl form"  
FT Modified-site 14 /note= "In amide form"  
FT  
FT  
XX  
XX W09640189-A1.  
PN  
XX 19-DEC-1996.  
PD  
XX 05-JUN-1996; 96WO-US008998.  
PF  
XX 07-JUN-1995; 95US-00472371.  
PR  
XX 07-JUN-1995; 95US-004723604.  
PR  
XX 07-JUN-1995; 95US-00476168.  
PR  
XX 07-JUN-1995; 95US-00478128.  
PR  
XX 07-JUN-1995; 95US-00484090.  
PR  
XX 07-JUN-1995; 95US-00485301.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
XX Dower WJ, Barrett RM, Cwirja SE, Duffin DJ, Gates CM, Johnson SS;  
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
PI  
XX WPI; 1997-051883/05.  
DR  
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide  
XX mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.  
PT  
XX  
PS Claim 30; Page 91; 106pp; English.  
PS  
XX The present sequence is a compound which binds to thrombopoietin (TPO)  
CC receptor (TR). The compound can be used for treating patients suffering  
CC from haematological disorders and thrombocytopenia resulting from  
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide  
CC may also be used to maintain the proliferation and growth of TPO-  
CC dependent cell lines and for use in biological research, for detecting  
XX TPO receptors on living cells  
XX  
XX Sequence 14 AA;  
Query Match 74.6%; Score 85; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 CADGPTLRWISFC 16  
DB 1 CADGPTLRWISFC 14

ID	AA	standard; protein; 14 AA.
XX	AA	09462;
XX	DT	10-SEP-1997 (first entry)
XX	DE	Thrombopoietin receptor binding compound peptide.
XX	KW	Haematology; thrombocytopenia; TPO; TR; proliferation;
XX	KW	bone marrow transfusion; chemotherapy; radiation therapy.
XX	OS	Synthetic.
XX	PH	Key
XX	FT	Misc-difference
XX	FT	1.14
XX	FT	/note= "preferably linkages are selected from: -
XX	FT	CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
XX	FT	; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
XX	FT	lower alkyl"
XX	FT	1
XX	FT	/note= "preferably N-terminus is selected from: -NR1; -
XX	FT	NHC(O)R; -NHC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
XX	FT	benzyloxycarbonyl-NH; benzyloxycarbonyl-NH with 1-3
XX	FT	substitutions on the phenyl ring selected from lower
XX	FT	alkyl, lower alkoxy, chloro, bromo; where R and R1 are
XX	FT	independently selected from hydrogen and lower alkyl"
XX	FT	14
XX	FT	/note= "preferably C-terminus is -C(O)R2 where R2 is
XX	FT	selected from hydroxy, lower alkoxy, and -NR3R4, where R3
XX	FT	and R4 are independently selected from hydrogen and lower
XX	FT	alkyl, and where the nitrogen atom of the -NR3R4 group
XX	FT	can optionally be the amine group of the N-terminus of
XX	FT	the peptide forming a cyclic peptide"
XX	PN	WO9640189-A1.
XX	PD	19-DEC-1996.
XX	PF	05-JUN-1996; 96WO-US008998.
XX	PR	07-JUN-1995; 95US-00472371.
XX	PR	07-JUN-1995; 95US-00473604.
XX	PR	07-JUN-1995; 95US-00476168.
XX	PR	07-JUN-1995; 95US-00478128.
XX	PR	07-JUN-1995; 95US-00484090.
XX	PR	07-JUN-1995; 95US-00485301.
XX	PA	(GLAXO) GLAXO GROUP LTD.
XX	PI	Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX	PI	Mattheakis LC, Schatz PJ, Wagstroom CR, Wrighton NC;
XX	XX	WPI, 1997-051883/05.
XX	DR	Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX	PT	mimetic(s) - useful in treatment of haematological disorders, esp.
XX	PT	thrombocytopenia resulting from chemotherapy, etc.
XX	PS	Claim 18; Page 89; 106dp; English.
XX	XX	The present sequence is a compound which binds to thrombopoietin (TPO)
XX	XX	receptor (TR). It has a molecular weight of < 8000 Da, and a binding
XX	XX	affinity to TR as expressed by an IC50 of no more than about 100 mM. The
XX	XX	compound (especially if modified, see features table) can be used for
XX	XX	treating patients suffering from haematological disorders and
XX	XX	thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX	XX	marrow transfusions. The peptide may also be used to maintain the
XX	XX	proliferation and growth of TPO-dependent cell lines and for use in
XX	XX	biological research, for detecting TPO receptors on living cells

[illegible]

## RESULT 15

AAW09482 AAW09482 standard; protein; 14 AA.

XX AAW09482;

DT 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding peptide.

DE Haematology; thrombocytopenia; TPO; TR; proliferation;

KM bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX WO9640189-A1.

PD 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

XX Thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 26; 106pp; English.

XX The present sequence is a peptide which binds to thrombopoietin (TPO)

XX receptor (TR). The compound can be used for treating patients suffering

XX from haematological disorders and thrombocytopenia resulting from

XX chemotherapy, radiation therapy or bone marrow transfusions. The peptide

XX may also be used to maintain the proliferation and growth of TPO-

XX dependent cell lines and for use in biological research, for detecting

XX TPO receptors on living cells

XX Sequence 14 AA;

XX Query Match 74.6%; Score 85; DB 2; Length 14;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-05;

XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 3 CADGPTLRWISFC 16

XX DB 1 CADGPTLRWISFC 14

## RESULT 16

AAW33031 AAW33031 standard; peptide; 14 AA.

XX AAW33031;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

KM signal transduction; receptor activation; cell culture.

XX Synthetic.

XX Key Location/Qualifiers

XX Disulfide-bond 1. 14

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

XX Thrombopoietin receptor - useful in treatment of haematological

XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 30; Page 91; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a

XX molecular weight of less than 8000 Da and a TR binding affinity as

XX expressed by an IC50 of no more than about 100 microm. It can be used to

XX treat disorders which are susceptible to treatment with a thrombopoietin

XX agonist, preferably haematological disorders and thrombocytopenia

XX resulting from chemotherapy, radiation therapy or bone marrow

XX transfusions. It can also be used diagnostically, e.g. to investigate the

XX mechanism of thrombopoietin signal transduction and receptor activation,

XX or to maintain the proliferation and growth of thrombopoietin dependent

XX cell lines

XX Sequence 14 AA;

XX Query Match 74.6%; Score 85; DB 2; Length 14;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-05;

XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 3 CADGPTLRWISFC 16

XX DB 1 CADGPTLRWISFC 14

## RESULT 17

AAW36633 AAW36633 standard; peptide; 14 AA.

XX AAW36633;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

XX haematological disorder; thrombocytopenia; chemotherapy;

XX radiation therapy; bone marrow transfusion; diagnosis;

XX signal transduction; receptor activation; cell culture.

XX Synthetic.

XX WO9640750-A1.

XX 19-DEC-1996.

```

XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO ) GLAXO GROUP LTD.
PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 26; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX Sequence 14 AA;
SQ
Query Match 74.6%; Score 85; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 CADGPTLRWISFC 16
Db 1 CADGPTLRWISFC 14

```

RESULT 18  
AAW33029 standard; peptide; 14 AA.  
ID AAW33029 standard; peptide; 14 AA.  
AC AAW33029;  
XX  
XX 11-MAR-1998 (first entry)  
DT  
XX Thrombopoietin receptor binding peptide.  
DE  
XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
KW haematological disorder; thrombocytopenia; chemotherapy;  
KW radiation therapy; bone marrow transfusion; diagnosis;  
KW signal transduction; receptor activation; cell culture.  
XX  
XX Synthetic.  
OS  
XX WO9640750-A1.  
XX  
XX 19-DEC-1996.  
PD  
XX 07-JUN-1996; 96WO-US009623.  
PF  
XX 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
XX  
XX (GLAXO ) GLAXO GROUP LTD.  
XX  
XX Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX WPI; 1997-052226/05.  
XX  
XX Peptides and peptide mimetics which bind to and activate the

```

PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 19; Page 89; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
XX Sequence 14 AA;
SQ
Query Match 74.6%; Score 85; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 CADGPTLRWISFC 16
Db 1 CADGPTLRWISFC 14

```

RESULT 19  
AAW35401 standard; peptide; 14 AA.  
ID AAW35401 standard; peptide; 14 AA.  
AC AAW35401;  
XX  
XX 11-MAR-1998 (first entry)  
DT  
XX Thrombopoietin receptor binding peptide.  
DE  
XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
KW haematological disorder; thrombocytopenia; chemotherapy;  
KW radiation therapy; bone marrow transfusion; diagnosis;  
KW signal transduction; receptor activation; cell culture.  
XX  
XX Synthetic.  
OS  
XX Key Location/Qualifiers  
FH Disulfide-bond 1. .14  
FT Modified-site 14  
FT /note= "NH2-D-Cys"  
XX  
XX WO9640750-A1.  
XX  
XX 19-DEC-1996.  
PD  
XX 07-JUN-1996; 96WO-US009623.  
PF  
XX 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
XX  
XX (GLAXO ) GLAXO GROUP LTD.  
XX  
XX Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX WPI; 1997-052226/05.  
XX  
XX Peptides and peptide mimetics which bind to and activate the  
PT thrombopoietin receptor - useful in treatment of haematological  
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
XX  
XX Example 6; Page 63; 106pp; English.  
XX  
XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX Sequence 14 AA;  
 SQ

Query Match 74.6%; Score 85; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14

RESULT 20

AAW36647  
 ID AAW36647 standard; peptide; 14 AA.

XX AAW36647;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

XX haematological disorder; thrombocytopenia; chemotherapy;

XX radiation therapy; bone marrow transfusion; diagnosis;

XX signal transduction; receptor activation; cell culture.

XX Synthetic.

XX W09640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

XX Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

XX thrombopoietin receptor - useful in treatment of haematological

XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 26; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

XX used to treat disorders which are susceptible to treatment with a

XX thrombopoietin agonist, preferably haematological disorders and

XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone

XX marrow transfusions. It can also be used diagnostically, e.g. to

XX investigate the mechanism of thrombopoietin signal transduction and

XX receptor activation, or to maintain the proliferation and growth of

Db |||||  
 1 CADGPTLRWISFC 14

RESULT 21

AAW35400  
 ID AAW35400 standard; peptide; 14 AA.

XX AAW35400;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

XX haematological disorder; thrombocytopenia; chemotherapy;

XX radiation therapy; bone marrow transfusion; diagnosis;

XX signal transduction; receptor activation; cell culture.

XX Synthetic.

XX W09640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

XX Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

XX thrombopoietin receptor - useful in treatment of haematological

XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 6; Page 63; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

XX used to treat disorders which are susceptible to treatment with a

XX thrombopoietin agonist, preferably haematological disorders and

XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone

XX marrow transfusions. It can also be used diagnostically, e.g. to

XX investigate the mechanism of thrombopoietin signal transduction and

XX receptor activation, or to maintain the proliferation and growth of

XX thrombopoietin dependent cell lines

XX Sequence 14 AA;

Query Match 74.6%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFC 16

Db 1 CADGPTLRWISFC 14

RESULT 22

AAW33032  
 ID AAW33032 standard; peptide; 14 AA.

```

XX AC AAW3032;
XX KW 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
XX KW haematological disorder; thrombocytopaenia; chemotherapy;
XX KW radiation therapy; bone marrow transfusion; diagnosis;
XX KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Disulfide-bond 1.14
XX FT Modified-site /note="acylated"
XX FT Modified-site 14
XX FT /note="amidated"
XX PN MO9640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAXO ) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX PT Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX CC Claim 30; Page 91; 106pp; English.
XX CS The present peptide binds the thrombopoietin receptor (TR), has a
XX CC molecular weight of less than 8000 Da and a TR binding affinity as
XX CC expressed by an IC50 of no more than about 100 microm. It can be used to
XX CC treat disorders which are susceptible to treatment with a thrombopoietin
XX CC agonist, preferably haematological disorders and thrombocytopaenia
XX CC resulting from chemotherapy, radiation therapy or bone marrow
XX CC transfusions. It can also be used diagnostically, e.g. to investigate the
XX CC mechanism of thrombopoietin signal transduction and receptor activation,
XX CC or to maintain the proliferation and growth of thrombopoietin dependent
XX CC cell lines
XX SQ Sequence 14 AA;
XX QY Query Match 74.6%; Score 85; DB 2; Length 14;
XX DB Best Local Similarity 100.0%; Pred. No. 1.7e-05;
XX DB Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 3 CADGPTLREWISFC 16
XX DB 1 CADGPTLREWISFC 14
XX RESULT 23
XX ID AAB17014 standard; peptide; 14 AA.
XX AC AAB17014;
XX DE 31-OCT-2000 (first entry)
XX KW

```

```

DE TPO-mimetic peptide sequence SEQ ID NO:70.
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX KW autoimmune disease; cytostatic; antiaesthetic; thrombolytic; VEGF;
XX KW immunosuppressive; EPO; TPO; C14A4; mimetic; IL-1; TNF; antagonist; MMP;
XX KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX OS Synthetic.
XX PN WO200024782-A2.
XX PD 04-MAY-2000.
XX PF 25-OCT-1999; 99WO-US025044.
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX PA (AMGEN-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham J, Boone TC;
XX DR WPI; 2000-350702/30.
XX PT Novel composition of matter comprising an Fc domain and pharmacologically
XX PT active peptides, useful for treating cancer and autoimmune diseases.
XX PS Claim 19; Page 218; 608pp; English.
XX CS The present invention describes composition of matter (I) comprising an
XX CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
XX CC independently selected from -(L1)-c-P1, -(L1)c-P1-(L2)-d-P2, -(L1)c-P1-
XX CC (L2)-d-P2-(L3)-e-P3, or -(L1)c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,
XX CC P3, and P4 = are each independently sequences of pharmacologically active
XX CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX CC of a and b is 1. The composition can have cytostatic, antiaesthetic,
XX CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX CC cells from the present invention can be used for producing pharmaceutical
XX CC compositions. The compositions are useful for treating cancer, asthma,
XX CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX CC a Fab domain) can provide a longer half-life or incorporate functions, and
XX CC such as Fc receptor binding, protein A binding, complement fixation, and
XX CC possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to
XX CC AAB18003 represent nucleotide and amino acid sequences used in the
XX CC exemplification of the present invention
XX SQ Sequence 14 AA;
XX QY Query Match 74.6%; Score 85; DB 3; Length 14;
XX DB Best Local Similarity 100.0%; Pred. No. 1.7e-05;
XX DB Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 3 CADGPTLREWISFC 16
XX DB 1 CADGPTLREWISFC 14
XX RESULT 24
XX ID AAU25826 standard; peptide; 14 AA.
XX AC AAU25826;
XX DE 17-DEC-2001 (first entry)
XX KW Human thrombopoietin receptor (TPO-R) activator peptide #12.
XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

```

KM haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KM bone marrow transplantation; hematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 PE  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 67-68; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 74.6%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14  
 RESULT 25  
 AAU25852  
 ID AAU25852 standard; peptide; 14 AA.  
 XX  
 AC AAU25852;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #38.

KM Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KM haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KM bone marrow transplantation; hematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 PE  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 74.6%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14  
 RESULT 26  
 AAU25866  
 ID AAU25866 standard; peptide; 14 AA.  
 XX  
 AC AAU25866;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #52.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iacti gene.  
 XX Homo sapiens.  
 XX US6251864-B1.  
 XX 26-JUN-2001.  
 XX 01-MAR-2000; 2000US-00516704.  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX (GLAXO ) GLAXO GROUP LTD.  
 XX Dower WJ, Barrett RM, Cwiria SE, Gates CM, Scharz RJ,  
 PI Balasubramanian P, Weststrom CR, Hendren RM, Deprince RB, Poddaturi S,  
 PI Yin Q.  
 XX WPI; 2001-564142/63.  
 XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX Disclosure; Col 20; 128pp; English.  
 XX Sequences ANU25815-ANU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX Sequence 14 AA;  
 SQ  
 Query Match 74.6%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 RESULT 27  
 ABB72900  
 ID ABB72900 standard; peptide; 14 AA.  
 XX  
 AC ABB72900;  
 XX  
 DT 05-APR-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:70.  
 XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
 XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumor; immunosuppressive;  
 KW cytostatic; antineumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianemic; anorectic; antifertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200183525-A2.  
 XX 08-NOV-2001.  
 PD 02-MAY-2001; 2001WO-US014310.  
 XX 03-MAY-2000; 2000US-00563286.  
 XX (AMGE-) AMGEN INC.  
 XX Feige U, Liu C, Cheatham JC, Boone TC, Gudae JW,  
 PI WPI; 2002-130313/17.  
 XX Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX Claim 39; Page 44; 176pp; English.  
 PS The present invention describes a vehicle-peptide molecule (I) or its  
 XX multimers. (I) can have antiinflammatory, antitumor, immunosuppressive,  
 CC cyostatic, antineumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (II) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX Sequence 14 AA;  
 SQ  
 Query Match 74.6%; Score 85; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 RESULT 28



ADJ73051 standard; peptide; 14 AA.  
 ID ADJ73051  
 XX  
 AC ADJ73051;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE TPO mimetic peptide sequence SegID 505.  
 XX  
 KM mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KM immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
 KM TPO.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003084477-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 24-MAR-2003; 2003WO-US009139.  
 XX  
 PR 29-MAR-2002; 2002US-0368791P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;  
 XX  
 DR WPI; 2003-804237/75.  
 XX  
 PT New CDR mimetibody comprising a portion of a heavy or light chain  
 PT variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PT cardiovascular or neurologic disease.  
 XX  
 PS Disclosure; SEQ ID NO 505; 97pp; English.  
 XX  
 CC This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LBR). The present invention describes human  
 CC mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neurologic  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 74.6%; Score 85; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 XX  
 RESULT 29  
 ADJ52686 standard; peptide; 14 AA.  
 ID ADJ52686  
 XX  
 AC ADJ52686;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX

DE CHI deleted mimetibody-related peptide SegID505.  
 XX  
 KM CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 KM hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KM fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KM arrhythmia; hypertension; heart failure; neurodegenerative;  
 KM multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KM cancerous condition; infectious disease; bacterial infection;  
 KM viral infection; fungal infection.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 XX  
 PN WO2004002417-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 27-JUN-2003; 2003WO-US020347.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nespor TC;  
 PI Kutolowski KA;  
 XX  
 DR WPI; 2004-082870/08.  
 XX  
 PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 XX  
 PS Claim 2; SEQ ID NO 505; 129pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 74.6%; Score 85; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 DB 1 CADGPTLRWISFC 14  
 XX  
 RESULT 30  
 ADJ51647 standard; peptide; 14 AA.  
 ID ADJ51647  
 XX  
 AC ADJ51647;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CHI deleted mimetibody-related peptide SegID505.  
 XX  
 KM CHI deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
 KW ophthalmological; nephrotoxic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstetric disorder; haematological disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 XX Unidentified.  
 OS Synthetic.  
 XX  
 PN WO2004002424-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heaver GA, Knight DM, Ghayeb J, Scallion BJ, Nespor TC;  
 PI Kutolowski KA;  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 15; SEQ ID NO 505; 123pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotoxic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstructive, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 74.6%; Score 85; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 CADGPTLRWISFC 16  
 Db 1 CADGPTLRWISFC 14

ID AAM09467 standard; protein; 13 AA.  
 XX  
 AC AAM09467;  
 XX  
 DT 10-SEP-1997 (first entry)  
 XX  
 DE Thrombopoietin receptor binding compound cyclic peptide.  
 KW Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;  
 KW bone marrow transfusion; chemotherapy; radiation therapy.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "The Ala is linked with the modified Cys at  
 FT position 13"  
 FT Modified-site 14  
 FT /label= OTHER  
 FT /note= "S-carboxymethyl-L-cysteine alpha-carboxamide;  
 FT forming a linkage onto the Ala at position one with the  
 FT delta C of this residue"  
 XX  
 PN WO9640189-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 05-JUN-1996; 96WO-US008998.  
 XX  
 PR 07-JUN-1995; 95US-00472371.  
 PR 07-JUN-1995; 95US-00473604.  
 PR 07-JUN-1995; 95US-00476168.  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00484090.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 DR WPI; 1997-051883/05.  
 XX  
 PT Thrombopoietin receptor-binding/activating peptide(s) and peptide  
 PT mimetic(s) - useful in treatment of haematological disorders, esp.  
 PT thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Claim 30; Page 91; 106pp; English.  
 XX  
 CC The present sequence is a compound which binds to thrombopoietin (TPO)  
 CC receptor (TR). The compound can be used for treating patients suffering  
 CC from haematological disorders and thrombocytopenia resulting from  
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide  
 CC may also be used to maintain the proliferation and growth of TPO-  
 CC dependent cell lines and for use in biological research, for detecting  
 CC TPO receptors on living cells  
 XX  
 SQ Sequence 13 AA;  
 Query Match 66.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 Db 1 ADGPTLRWISFC 13

```

XX 11-MAR-1998 (first entry)
DT Thrombopoietin receptor binding peptide.
DE
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "COCH2-alanine linked via CH2 group to Cys13"
FT Modified-site 13 /note= "NH2-cytosine linked via sulphoxidised thiol group to Ala1"
FT
FT
XX WO9640750-A1.
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Example 6; Page 63; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX Sequence 13 AA;
SQ
Query Match 66.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ADGPTLRWISFC 16
DB 1 ADGPTLRWISFC 13

```

```

KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical peptide"
FT Modified-site 13 /note= "NH2-Phe"
FT
XX WO9640750-A1.
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Example 9; Page 73; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX Sequence 13 AA;
SQ
Query Match 66.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CADGPTLRWISF 15
DB 1 CADGPTLRWISF 13

```

```

RESULT 33
AAW35417
ID AAW35417 standard; peptide; 13 AA.
XX
XX AAW35417;
AC
XX 11-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
DE
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;

```

```

RESULT 34
AAW33033
ID AAW33033 standard; peptide; 13 AA.
XX
XX AAW33033;
AC
XX 11-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
DE
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1

```

```

FT      /note= "COCH2-alanine linked via CH2 group to Cys13"
FT      13
FT      /note= "NH2-cytosine linked via thiol group to Ala1"
XX
XX      WO9640750-A1.
PN      19-DEC-1996.
XX
XX      07-JUN-1996;    96WO-US009623.
XX
XX      07-JUN-1995;    95US-00478128.
PR      07-JUN-1995;    95US-00485301.
XX
XX      (GLAX ) GLAXO GROUP LTD.
PI      Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI      Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX      WPI, 1997-052226/05.
DR
XX
XX      Peptides and peptide mimetics which bind to and activate the
PT      thrombopoietin receptor - useful in treatment of haematological
PT      disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX      Claim 30; Page 91; 106pp; English.
XX
XX      The present peptide binds the thrombopoietin receptor (TR), has a
CC      molecular weight of less than 8000 Da and a TR binding affinity as
CC      expressed by an IC50 of no more than about 100 microm. It can be used to
CC      treat disorders which are susceptible to treatment with a thrombopoietin
CC      agonist, preferably haematological disorders and thrombocytopenia
CC      resulting from chemotherapy, radiation therapy or bone marrow
CC      transfusions. It can also be used diagnostically, e.g. to investigate the
CC      mechanism of thrombopoietin signal transduction and receptor activation,
CC      or to maintain the proliferation and growth of thrombopoietin dependent
CC      cell lines
XX
XX      Sequence 13 AA;
SQ
XX

Query Match      66.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 ADGPTLRWISFC 16
      |||||
      1 ADGPTLRWISFC 13
DB

RESULT 35
AAW35413
ID      AAW35413 standard; peptide; 13 AA.
XX
XX      AAW35413;
AC
XX      11-MAR-1998 (first entry)
DT
XX
XX      Thrombopoietin receptor binding peptide.
DE
XX
XX      Thrombopoietin receptor; binding peptide; treatment; agonist;
KW      haematological disorder; thrombocytopenia; chemotherapy;
KW      radiation therapy; bone marrow transfusion; diagnosis;
KW      signal transduction; receptor activation; cell culture.
XX
XX      Synthetic.
OS
XX
XX      Key
FH      Modified-site      1 Location/Qualifiers
FT      /note= "Br-Ala"
FT      Modified-site      13
FT      /note= "NH2-Cys"
XX
XX      WO9640750-A1.

```

```

PD      19-DEC-1996.
XX
XX      07-JUN-1996;    96WO-US009623.
XX
XX      07-JUN-1995;    95US-00478128.
PR      07-JUN-1995;    95US-00485301.
XX
XX      (GLAX ) GLAXO GROUP LTD.
PI      Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI      Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX      WPI, 1997-052226/05.
DR
XX
XX      Peptides and peptide mimetics which bind to and activate the
PT      thrombopoietin receptor - useful in treatment of haematological
PT      disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX      Example 9; Page 73; 106pp; English.
XX
XX      The present peptide, which binds the thrombopoietin receptor (TR), can be
CC      used to treat disorders which are susceptible to treatment with a
CC      thrombopoietin agonist, preferably haematological disorders and
CC      thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC      marrow transfusions. It can also be used diagnostically, e.g. to
CC      investigate the mechanism of thrombopoietin signal transduction and
CC      receptor activation, or to maintain the proliferation and growth of
CC      thrombopoietin dependent cell lines
XX
XX      Sequence 13 AA;
SQ
XX

Query Match      66.7%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 ADGPTLRWISFC 16
      |||||
      1 ADGPTLRWISFC 13
DB

RESULT 36
AAW35406
ID      AAW35406 standard; peptide; 13 AA.
XX
XX      AAW35406;
AC
XX      11-MAR-1998 (first entry)
DT
XX
XX      Thrombopoietin receptor binding peptide.
DE
XX
XX      Thrombopoietin receptor; binding peptide; treatment; agonist;
KW      haematological disorder; thrombocytopenia; chemotherapy;
KW      radiation therapy; bone marrow transfusion; diagnosis;
KW      signal transduction; receptor activation; cell culture.
XX
XX      Synthetic.
OS
XX
XX      Key
FH      Modified-site      1 Location/Qualifiers
FT      /note= "CO-CH (Ph)-alanine linked via CH group to Cys13"
FT      Modified-site      13
FT      /note= "NH2-cytosine linked via thiol group to Ala1"
XX
XX      WO9640750-A1.
XX
XX      19-DEC-1996.
PD
XX
XX      07-JUN-1996;    96WO-US009623.
XX
XX      07-JUN-1995;    95US-00478128.
PR      07-JUN-1995;    95US-00485301.
XX
XX      (GLAX ) GLAXO GROUP LTD.

```

XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 6; Page 64; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 SQ Sequence 13 AA;  
 Query Match 66.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00033; Mismatches 0; Gaps 0;  
 Matches 13; Conservative 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13  
 Db  
 RESULT 37  
 AAW35422  
 ID AAW35422 standard; peptide; 13 AA.  
 XX  
 AC AAW35422;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "optionally acylated"  
 FT Cross-links 13 /note= "linked via disulfide bond to Cys13 of identical  
 FT peptide"  
 FT  
 XX WO9640750-A1.  
 PN  
 XX 19-DEC-1996.  
 PD  
 XX 07-JUN-1996; 96WO-US009623.  
 PF  
 XX 07-JUN-1995; 95US-00478128.  
 PR  
 XX 07-JUN-1995; 95US-00485301.  
 PA  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 XX Example 9; Page 74; 106pp; English.  
 PS  
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 SQ Sequence 13 AA;  
 Query Match 66.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00033; Mismatches 0; Gaps 0;  
 Matches 13; Conservative 0; Indels 0; Gaps 0;  
 QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13  
 Db  
 RESULT 38  
 AAW35397  
 ID AAW35397 standard; peptide; 13 AA.  
 XX  
 AC AAW35397;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "COCH2-alanine linked via CH2 group to Cys13"  
 FT Modified-site 13 /note= "NH2-cytosine linked via thiol group to Ala1"  
 FT  
 XX WO9640750-A1.  
 PN  
 XX 19-DEC-1996.  
 PD  
 XX 07-JUN-1996; 96WO-US009623.  
 PF  
 XX 07-JUN-1995; 95US-00478128.  
 PR  
 XX 07-JUN-1995; 95US-00485301.  
 PA  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 6; Page 63; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transplants. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX Sequence 13 AA;

Query Match 66.7%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ADGPTLRWISFC 16  
 |||||  
 1 ADGPTLRWISFC 13

Db

RESULT 39

AAU25997

ID AAU25997 standard; peptide; 13 AA.

XX AAU25997;

DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #183.

DE Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

OS US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;  
 PI Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 143-144; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 13 AA;

Query Match 66.7%; Score 76; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CADGPTLRWISF 15  
 |||||  
 1 CADGPTLRWISF 13

Db

RESULT 40

AAU25984

ID AAU25984 standard; peptide; 13 AA.

XX AAU25984;

DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #170.

DE Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

OS US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;  
 PI Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 137; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO

CC thrombopoietin dependent cell lines

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ADGPTLREWISFC 16  
 |||||  
 XX  
 Db 2 ADGPTLREWISFC 14

## RESULT 43

AAW35402  
 ID AAW35402 standard; peptide; 14 AA.

AC AAW35402;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1. .14

FT Modified-site /note= "D-form residue, Penicillamine"

FT Modified-site 14 /note= "NH2-D-Cys"

XX MO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Example 6; Page 64; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC marrow transplants. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX Sequence 14 AA;

QY Query Match 66.7%; Score 76; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 ADGPTLREWISFC 16

2 ADGPTLREWISFC 14

ID AAU25987 standard; peptide; 14 AA.

XX AAU25987;

XX 18-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #173.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

KW bone marrow transplantation; haematological disorder; platelet disorder;

KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;

KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;

KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Depriente RB, Podduturi S;

XX yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 139; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

CC progenitors alone or in conjunction with additional cytokines

XX Sequence 14 AA;

QY Query Match 66.7%; Score 76; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 CADGPTLREWISF 15

1 CADGPTLREWISF 13

RESULT 45



Search completed: September 1, 2005, 16:12:10  
Job time : 87.3453 secs

```

AAU25983
ID AAU25983 standard; peptide: 14 AA.
XX
AC AAU25983;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #169.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; Iaci gene.
XX
OS Homo sapiens.
XX
PN US651864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAXO ) GLAXO GROUP LTD.
XX
PI Dower WJ, Barreclt RM, Cwiria SE, Gates CM, Schatz PJ,
PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 135-137; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SO Sequence 14 AA;

Query Match 56.7%; Score 76; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 ADGPTLRWISFC 16
|||||
2 ADGPTLRWISFC 14

```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: September 1, 2005, 15:57:33 ; Search time 14.4892 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-8

Perfect score: 114  
Sequence: 1 GGCADGPTLRWISFCGSK 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	43.0	245	2	T47701
2	47	41.2	475	2	T33943
3	46	40.4	346	2	A58583
4	45	39.5	108	2	T49731
5	45	39.5	180	2	T49494
6	45	39.5	421	2	T22969
7	45	39.5	499	2	S51089
8	44	38.6	346	2	T19008
9	44	38.6	371	2	D75266
10	44	38.6	490	2	T08084
11	44	38.6	526	2	A86440
12	44	38.6	974	2	S34189
13	44	38.6	1022	1	S00503
14	44	38.6	1023	2	A24414
15	44	38.6	376	2	T39685
16	43.5	38.2	1499	2	A88813
17	43	37.7	113	2	D72595
18	43	37.7	115	2	T15386
19	43	37.7	192	1	A24902
20	43	37.7	230	2	A48685
21	43	37.7	233	2	A82768
22	43	37.7	246	2	T15988
23	43	37.7	268	2	D97548
24	43	37.7	276	2	A38654
25	43	37.7	434	2	S21324
26	43	37.7	629	2	A82497
27	43	37.7	953	2	S54478
28	43	37.7	1010	2	B37227
29	43	37.7	1010	2	B37227

30	43	37.7	1013	1	S00801	Na+/K+-exchanging
31	43	37.7	1013	2	C24639	Na+/K+-exchanging
32	43	37.7	1017	2	A57227	Na+/K+-exchanging
33	43	37.7	1020	2	A34474	Na+/K+-exchanging
34	43	37.7	1020	2	B24639	Na+/K+-exchanging
35	43	37.7	1021	1	PMSHNA	Na+/K+-exchanging
36	43	37.7	1021	1	S04630	Na+/K+-exchanging
37	43	37.7	1021	2	A28159	Na+/K+-exchanging
38	43	37.7	1021	2	B24862	Na+/K+-exchanging
39	43	37.7	1022	2	S49127	Na+/K+-exchanging
40	43	37.7	1023	1	A24639	Na+/K+-exchanging
41	43	37.7	1023	1	S24650	Na+/K+-exchanging
42	43	37.7	1025	2	A60444	Na+/K+-exchanging
43	43	37.7	1027	1	PMCMNA	Na+/K+-exchanging
44	43	37.7	1038	1	S03632	Na+/K+-exchanging
45	42.5	37.3	210	2	A42687	neurotrophin-4 pre
46	42.5	37.3	353	2	T32638	hypothetical prote
47	42.5	37.3	1004	2	JH0470	Na+/K+-exchanging
48	42.5	37.3	1302	2	T00038	hypothetical prote
49	42	36.8	98	2	A70301	ribosomal protein
50	42	36.8	141	2	AH2829	conserved hypothet
51	42	36.8	141	2	F97607	hypothetical prote
52	42	36.8	192	1	S28148	erythropoietin pre
53	42	36.8	312	2	F86876	hypothetical prote
54	42	36.8	440	2	F81555	glutamate-1-semial
55	42	36.8	440	2	B86508	glutamate-1-semial
56	42	36.8	440	2	G72114	glutamate-1-semial
57	42	36.8	473	2	T31717	hypothetical prote
58	42	36.8	522	2	D69226	hypothetical prote
59	42	36.8	522	2	S62941	probable membrane
60	42	36.8	725	2	A11544	conserved hypothet
61	42	36.8	816	2	T08978	serine proteinase
62	42	36.8	842	2	T12091	serine phosphoryla
63	41.5	36.4	108	2	G82991	thioredoxin P45240
64	41.5	36.4	41.5	1	G45522	similar to gibbera
65	41.5	36.4	209	2	B42687	neurotrophin-4 pre
66	41.5	36.4	359	2	T15470	hypothetical prote
67	41	36.0	132	1	G69256	conserved hypothet
68	41	36.0	189	2	C71943	hypothetical prote
69	41	36.0	189	2	S07755	hypothetical prote
70	41	36.0	206	2	T22345	hypothetical prote
71	41	36.0	239	2	AC2745	glycerophosphoryl
72	41	36.0	245	2	JC7273	inducible mast cel
73	41	36.0	246	2	B97526	hypothetical prote
74	41	36.0	273	2	H70849	hypothetical prote
75	41	36.0	274	2	A45754	tryptase (BC 3.4.2
76	41	36.0	275	2	C35863	hypothetical prote
77	41	36.0	298	2	T23362	hypothetical prote
78	41	36.0	410	1	DEPSXA	2-oxoisovalerate d
79	41	36.0	410	2	C83365	hypothetical prote
80	41	36.0	473	2	B84853	conserved hypothet
81	41	36.0	494	2	H82489	protein R09B5.11 l
82	41	36.0	576	2	C88950	coagulation factor
83	41	36.0	593	2	S45281	hypothetical prote
84	41	36.0	618	2	T48193	hypothetical prote
85	41	36.0	916	2	H72372	exonuclease ABC c
86	41	36.0	929	2	S75098	serine phosphoryla
87	41	36.0	955	2	T10947	serine phosphoryla
88	41	36.0	966	1	PHOAG	search phosphoryla
89	41	36.0	971	2	T09210	search phosphoryla
90	41	36.0	1000	2	S47243	protein P27F5.11 l
91	41	36.0	1313	2	B96509	hypothetical prote
92	41	36.0	1522	2	C96578	hypothetical prote
93	41	36.0	1616	2	T17884	S-layer protein -
94	40.5	35.5	369	2	B84542	hypothetical prote
95	40.5	35.5	1101	2	T16840	hypothetical prote
96	40.5	35.5	1363	2	T43220	insulin-like growt
97	40	35.1	152	2	S21826	T-cell receptor be
98	40	35.1	155	2	S23629	hypothetical prote
99	40	35.1	157	2	B83066	hypothetical prote
100	40	35.1	169	1	ICMS2	interleukin-2 prec

## ALIGNMENTS

## RESULT 1

T47701 translation initiation factor eIF-6-like protein [imported] - Arabidopsis thaliana

N/Alternate names: protein F1116.30

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C/Accession: T47701

R/Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.

submitted to the Protein Sequence Database, March 2000

A/Reference number: Z24473

A/Accession: T47701

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-245 &lt;BEN&gt;

A/Cross-references: UNIPROT:Q9M060; EMBL:AL161667

A/Experimental source: cultivar Columbia; BAC clone F1116

C/Genetics:

A/Map position: 3

A/Insertions: 4/1; 36/2; 65/1; 80/1; 123/3; 160/3

C/Superfamily: conserved hypothetical protein YPR016c

Query Match 43.0%; Score 49; DB 2; Length 245;

Best Local Similarity 57.1%; Pred. No. 9.2;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 ADGPTLRWISFCG 17

DB 194 AAGTVDWTSFCG 207

## RESULT 2

T33943 hypothetical protein C01B4.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T33943

R/Smith, A.; Mameley, P.; Fromack, W.

submitted to the EMBL Data Library, February 1999

A/Description: The sequence of C. elegans cosmid C01B4.

A/Reference number: Z21443

A/Accession: T33943

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-475 &lt;SMI&gt;

A/Cross-references: UNIPROT:Q9UAT5; EMBL:AF125952; PIDN:AA014699.1; GSPDB:GN00023; CESP:

A/Experimental source: strain Bristol N2; clone C01B4

C/Genetics:

A/Map position: 5

A/Insertions: 45/2; 80/1; 118/2; 189/3; 239/2; 340/3; 433/3

Query Match 41.2%; Score 47; DB 2; Length 475;

Best Local Similarity 50.0%; Pred. No. 33;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCG 18

DB 268 CTDRCVLSAWVSFLGG 283

## RESULT 3

A58583 testosterone-resistant immunity-associated protein IAP38 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 25-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004

C/Accession: A58583

R/Kneucken, J.; Schmitt-Wrede, H.P.; Markmann-Mulisch, U.; Wunderlich, F.

Biochem. Biophys. Res. Commun. 230, 167-170, 1997

A/Title: Novel gene expressed in spleen cells mediating acquired testosterone-resistant

A/Reference number: A58583; MUID:97148595; PMID:9020038

A/Accession: A58583

A/Molecule type: mRNA

A/Residues: 1-346 &lt;KRU&gt;

A/Cross-references: UNIPROT:P70224; GB:Y08026; NID:g1550784; PIDN:CAA69283.1; PID:g1550

C/Comment: This protein is a plasma membrane protein with two membrane-spanning domains

C/Accession: T47701

C/Genetics:

A/Map position: 3

F/148-167/Domain: transmembrane #status predicted &lt;TW1&gt;

F/320-335/Domain: transmembrane #status predicted &lt;TW2&gt;

Query Match 40.4%; Score 46; DB 2; Length 346;

Best Local Similarity 41.2%; Pred. No. 35;

Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCGK 19

DB 213 CTDRALRDLVVAECGR 229

## RESULT 4

T49731 hypothetical protein B24B19.30 [imported] - Neurospora crassa

C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000

C/Accession: T49731

R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura

submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25022

A/Accession: T49731

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-108 &lt;SCH&gt;

A/Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.30

A/Experimental source: BAC clone B24B19; strain OR74A

C/Genetics:

A/Map position: 6

C/Superfamily: Neurospora crassa hypothetical protein B24B19.30

Query Match 39.5%; Score 45; DB 2; Length 108;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFC 16

DB 70 CCGQPLRMLSWC 83

## RESULT 5

T44944 hypothetical protein 5 [imported] - Natronobacterium pharaonis

C/Species: Natronobacterium pharaonis

C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C/Accession: T44944

R/Mattar, S.; Engelhard, W.

Eur. J. Biochem. 250, 332-341, 1997

A/Title: Cytochrome b23 from Natronobacterium pharaonis: An archaeal four-subunit cyto

A/Reference number: Z22876; MUID:98088958; PMID:9428682

A/Accession: T44944

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-180 &lt;MAT&gt;

A/Cross-references: UNIPROT:Q07291; EMBL:Y10500; PIDN:CAA71527.1

A/Experimental source: strain Sp1/28

C/Genetics:

C/Superfamily: conserved hypothetical protein AF1745

Query Match 39.5%; Score 45; DB 2; Length 180;

Best Local Similarity 77.8%; Pred. No. 27;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 9 LREWISFCG 17  
 Db 116 LLEWISFCG 124

## RESULT 6

T22969  
 hypothetical protein F59A1.13 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T22969

R:Mottimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19644

A:Accession: T22969

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-421 <WIL>

A:Cross-references: UNIPROT:O9XUV7; EMBL:Z81557; PIDN:CA804538.1; GSPDB:GN00023; CESP:FS

A:Experimental source: clone F59A1

C:Genetics:

A:Gene: CESP:F59A1.13

A:Map position: 5

A:introns: 27/1; 116/1; 245/3; 286/3; 340/3; 381/3

Query Match

Best Local Similarity 39.5%; Score 45; DB 2; Length 421;  
 Best Local Similarity 50.0%; Pred. No. 58;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 3 CADGPTLREWISFCG 18

Db 214 CTDGTVLVGMISVFCG 229

## RESULT 7

S51089  
 ammonium transport protein MEP2 - yeast (Saccharomyces cerevisiae)

N:Alternate names: NH3 permease; protein JTA499; protein N1820; protein X

C:Species: Saccharomyces cerevisiae

C:Date: 10-May-1998 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004

C:Accession: S51089; S55142; S59247; S63087

R:Marini, A.M.; Andre, B.

submitted to the EMBL Data Library, December 1994

A:Reference number: S51089

A:Accession: S51089

A:Molecule type: DNA

A:Residues: 1-499 <MAR>

A:Cross-references: UNIPROT:P41948; EMBL:X83608; NID:G619513; PIDN:CA86884.1; PID:G6195

R:Mallet, L.; Bussereau, F.; Jacquet, M.

submitted to the EMBL Data Library, November 1994

A:Description: A 43.5 kb fragment of the chromosome XIV.

A:Reference number: S55136

A:Accession: S55142

A:Molecule type: DNA

A:Residues: 1-499 <MAL>

A:Cross-references: EMBL:Z46843; NID:G861113; PIDN:CA86884.1; PID:G854496

R:Mallet, L.; Bussereau, F.; Jacquet, M.

Yeast 11, 1195-1209, 1995

A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MPA2, MEP2, CAP/SRV2,

A:Accession: S59247

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-499 <MAW>

A:Cross-references: EMBL:Z46843; NID:G861113; PIDN:CA86884.1; PID:G854496

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994

R:Mallet, L.; Bussereau, F.; Jacquet, M.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63069

A:Accession: S63087

A:Molecule type: DNA

A:Residues: 1-499 <MAF>  
 A:Cross-references: EMBL:Z71418; NID:G1302090; PIDN:CA96025.1; PID:G1302091; MIPS:YNL1

A:Experimental source: strain S286C

C:Genetics:

A:Gene: SCD:MEP2

A:Cross-references: SGD:S0005086; MIPS:YNL142W

A:Map position: 14L

C:Function:

A:Description: ammonium transport

C:Superfamily: ammonium transport protein

C:Keywords: ammonium transport; transmembrane protein

F:35-51/Domain: transmembrane #status predicted <TM1>

F:62-78/Domain: transmembrane #status predicted <TM2>

F:123-139/Domain: transmembrane #status predicted <TM3>

F:154-170/Domain: transmembrane #status predicted <TM4>

F:228-244/Domain: transmembrane #status predicted <TM5>

F:288-304/Domain: transmembrane #status predicted <TM6>

F:306-322/Domain: transmembrane #status predicted <TM7>

F:397-413/Domain: transmembrane #status predicted <TM8>

Query Match

Best Local Similarity 39.5%; Score 45; DB 2; Length 499;  
 Best Local Similarity 38.5%; Pred. No. 68;

Matches 10; Conservative 1; Mismatches 7; Indels 8; Gaps 1;

Qy 1 GGCADGPTLREWISF-----CGG 18

Db 247 GGSAGNATIRAWYSIMSTNLAAACGG 272

## RESULT 8

T19008  
 hypothetical protein C06C6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T19008

R:McMurray, A.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19059

A:Accession: T19008

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-346 <WIL>

A:Cross-references: UNIPROT:O62030; EMBL:Z93374; PIDN:CA807554.1; GSPDB:GN00023; CESP:C

A:Experimental source: clone C06C6

C:Genetics:

A:Gene: CESP:C06C6.2

A:Map position: 5

A:introns: 109/1; 135/2; 160/2; 310/1

C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match

Best Local Similarity 38.6%; Score 44; DB 2; Length 346;  
 Best Local Similarity 56.2%; Pred. No. 68;

Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GCADGPTLREWISFCG 17

Db 183 GLADGSTITNDSPFG 198

## RESULT 9

D75266  
 cell division protein, FtsW/RodA/SpoVE family - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: D75266

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Frazer, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75266

A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-371 <WHI>  
A:Cross-references: UNIPROT:Q9R8J3; GB:AE002079; GB:AE000513; NID:G6460315; PIDN:AAE1203  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2497  
A:Map position: 1  
C:Superfamily: rod shape-determining protein

Query Match 38.6%; Score 44; DB 2; Length 371;  
Best Local Similarity 43.8%; Pred. No. 73;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCADGPTLRWISFC 17  
Db 77 GSGDPSGVRWRLSTAG 92

RESULT 10  
T09084  
phosphatidylinositol 3-kinase - Chlamydomonas reinhardtii (fragment)

C:Species: Chlamydomonas reinhardtii  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T09084  
R:Molendijk, A.J.; Irvine, R.F.  
Plant Mol. Biol. 37, 53-66, 1998  
A:Title: Inositolide signalling in Chlamydomonas: Characterization of a phosphatidylinositol  
A:Reference number: Z16411; MUID:98281574; PMID:9620264

A:Accession: T09084  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-490 <MOL>  
A:Cross-references: UNIPROT:O04270; EMBL:U97663; NID:G2109290; PIDN:AAC50018.1; PID:G210

A:Experimental source: strain CW-15  
C:Genetics:

A:Introns: 265/3; 331/3; 370/3; 455/1; 481/3

Query Match 38.6%; Score 44; DB 2; Length 490;  
Best Local Similarity 50.0%; Pred. No. 94;  
Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

QY 1 GGCA--DGPTLR--EWISFC 16  
Db 244 GGSSPGDGSYRWDEMLTFC 263

## RESULT 11

A86440

58.RK hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86440

R:Thelloglou, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A86440  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-526 <STO>

A:Cross-references: UNIPROT:Q9C868; GB:AB005172; NID:G11054679; PIDN:AA627899.1; GSPDB:C  
C:Genetics:  
A:Map position: 1

Query Match 38.6%; Score 44; DB 2; Length 526;  
Best Local Similarity 44.4%; Pred. No. 1e+02;

Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 GGCAAGPT--LRWISFC 16  
Db 395 GGRVGRGSPPLINQWIEFC 412

## RESULT 12

S34189

starch phosphorylase (EC 2.4.1.1) L - potato  
C:Species: Solanum tuberosum (potato)

C:Date: 03-Mar-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S53489; S34189

R:Somersdahl, U.; Baeser, A.; Greve, B.; Steup, M.  
Plant Mol. Biol. 27, 567-576, 1995

A:Title: A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inh  
A:Reference number: S53489; MUID:95201249; PMID:7894019

A:Accession: S53489  
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
A:Residues: 1-974 <SO2>

A:Cross-references: UNIPROT:P53535; EMBL:X73684; NID:G313348; PIDN:CAA52036.1; PID:G313  
C:Superfamily: glucan phosphorylase

C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphat  
F:820/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 38.6%; Score 44; DB 2; Length 974;  
Best Local Similarity 58.3%; Pred. No. 1.8e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DGPTLRWISFC 16  
Db 619 NGVTPRRWLSFC 630

## RESULT 13

S00503

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - Pacific electric ray  
C:Species: Torpedo californica (Pacific electric ray)

C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
C:Accession: S00503; S28885; S29880

R:Kakawani, K.; Noguchi, S.; Noda, M.; Takahashi, H.; Ohta, T.; Kawamura, M.; Nojima, H  
Nature 316, 733-736, 1985

A:Title: Primary structure of the alpha-subunit of Torpedo californica (Na(+)+K(+))ATPa  
A:Reference number: S00503; MUID:85296307; PMID:2893905

A:Accession: S00503  
A:Molecule type: mRNA

A:Residues: 1-1022 <KAW1>  
A:Cross-references: UNIPROT:P05025; EMBL:X02810; NID:G64399; PIDN:CAA26578.1; PID:G6440

A:Accession: S28885  
A:Molecule type: protein

A:Residues: 228-240/431-438/535-550/671-690/1011-1022 <KAW2>  
R:Ohta, T.; Nagano, K.; Yoshida, M.

Proc. Natl. Acad. Sci. U.S.A. 83, 2071-2075, 1986  
A:Title: The active site structure of Na(+)/K(+)-transporting ATPase: location of the 5

A:Reference number: S29880; MUID:86177549; PMID:3008150

A:Accession: S29880  
A:Molecule type: protein

A:Residues: 386-402/502-512/671-689/887-906 <OHT>  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C:Keywords: ATP; heterodimer; hydrolyase; ion transport; phosphoprotein; potassium trans  
F:96-120/Domain: transmembrane #status predicted <TM1>

F:110-149/Domain: transmembrane #status predicted <TM2>  
F:150-290/Domain: intracellular #status predicted <INT2>

F:291-313/Domain: transmembrane #status predicted <TM3>  
F:320-348/Domain: transmembrane #status predicted <TM4>

F:349-785/Domain: intracellular #status predicted <INT3>  
F:566-782/Domain: ATPase nucleotide-binding domain homology <ATN>

F:786-809/Domain: transmembrane #status predicted <TM5>  
F:848-873/Domain: transmembrane #status predicted <TM6>

F:874-951/Domain: intracellular #status predicted <INT4>  
F:952-977/Domain: transmembrane #status predicted <TM7>  
F:978-1022/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:507/Binding site: ATP (Lys) #status predicted  
 F:716,720,725/Active site: Asp, Asp, Lys #status predicted

Query Match 38.6%; Score 44; DB 1; Length 1022;  
 Best Local Similarity 70.0%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWSFC 16  
 Db 84 PTPPEWIKFC 93

## RESULT 14

A24414  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - human  
 N:Alternate names: sodium pump; sodium/potassium transporting ATPase alpha-A chain  
 C:Species: Homo sapiens (man)  
 C>Date: 02-Jun-1988 #sequence revision 02-Jun-1988 #text\_change 09-Jul-2004  
 C:Accession: A24414; A27795; A39910; I60116; S09171  
 R:Kawakami, K.; Ohta, T.; Nojima, H.; Nagano, K.  
 J. Biochem. 100, 389-397, 1986  
 A:Title: Primary structure of the alpha-subunit of human Na,K-ATPase deduced from cDNA  
 A:Reference number: A24414; MUID:87057096; PMID:2430951  
 A:Accession: A24414  
 A:Molecule type: mRNA  
 A:Residues: 1-1023 <KAW>  
 A:Cross-references: UNIPROT:P05023; EMBL:X04297; NID:g28926; PIDN:CAA27840.1; PID:g28927  
 R:Shull, M.M.; Lingrel, J.B.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
 A:Title: Multiple genes encode the human Na,K-ATPase catalytic subunit.  
 A:Reference number: A94158; MUID:87231946; PMID:3035563  
 A:Accession: A27795  
 A:Molecule type: DNA  
 A:Residues: 168-189;213-214,'X',216-244 <SHU>  
 R:Chehab, F.F.; Kan, Y.W.; Law, M.L.; Hartz, J.; Kao, F.T.; Blostein, R.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 7901-7905, 1987  
 A:Title: Human placental Na,K-ATPase alpha subunit: cDNA cloning, tissue expression, D  
 A:Reference number: A39910; MUID:88068506; PMID:2891135  
 A:Accession: A39910  
 A:Molecule type: preliminary  
 A:Status: preliminary  
 A:Residues: 199-942 <CHE>  
 A:Cross-references: GB:J03007  
 R:Shull, M.M.; Pugh, D.G.; Lingrel, J.B.  
 Genomics 6, 451-460, 1990  
 A:Title: The human Na,K-ATPase alpha 1 gene: characterization of the 5'-flanking region  
 A:Reference number: I60116; MUID:90228961; PMID:1970326  
 A:Accession: I60116  
 A:Status: translation not shown; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-61 <RES>  
 A:Cross-references: GB:M30310; NID:g179206; PIDN:AA51801.1; PID:g179208  
 C:Genetics:  
 A:Gene: GDB:ATP1A1  
 A:Cross-references: GDB:119711; OMTM:182310  
 A:Map position: 1p13-1p11  
 C:Superfamily: Na+/K+-translocating ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; heterodimer; hydrolase; ion transport; osmoregulation; phosphoprotein;  
 F:6-1022/Product: Na+/K+-translocating ATPase alpha-1 chain #status predicted <MAT>  
 F:6-95/Domain: Intracellular #status predicted <INT1>  
 F:96-130/Domain: transmembrane #status predicted <TM1>  
 F:130-149/Domain: transmembrane #status predicted <TM2>  
 F:150-290/Domain: intracellular #status predicted <INT2>  
 F:291-313/Domain: transmembrane #status predicted <TM3>  
 F:320-348/Domain: transmembrane #status predicted <TM4>  
 F:349-766/Domain: intracellular #status predicted <INT3>  
 F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:787-810/Domain: transmembrane #status predicted <TM5>  
 F:849-874/Domain: transmembrane #status predicted <TM6>  
 F:875-952/Domain: intracellular #status predicted <INT4>  
 F:953-978/Domain: transmembrane #status predicted <TM7>  
 F:979-1023/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:508/Binding site: ATP (Lys) #status predicted  
 F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 38.6%; Score 44; DB 2; Length 1023;  
 Best Local Similarity 70.0%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWSFC 16  
 Db 84 PTPPEWIKFC 93

## RESULT 15

T39685  
 conserved hypothetical protein SPBC1778.03c - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T39685  
 R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, March 1998  
 A:Reference number: Z21869  
 A:Accession: T39685  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-376 <OLI>  
 A:Cross-references: UNIPROT:Q9Y7J0; EMBL:AL049489; PIDN:CAB39798.1; GSPDB:GN00067; SPDB  
 A:Experimental source: strain 972h-; cosmid c1778  
 C:Genetics:  
 A:Gene: SPDB:SPBC1778.03c  
 A:Map position: 2  
 A:Introns: 11/2

Query Match 38.2%; Score 43.5; DB 2; Length 376;  
 Best Local Similarity 42.9%; Pred. No. 87;  
 Matches 9; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 1 GGCAGDGTLRBWS---FCGG 18  
 Db 164 GACAFARSIDWISRYRCPG 184

## RESULT 16

A89813  
 glutamate synthase large subunit [imported] - Staphylococcus aureus (strain N315).  
 C:Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence revision 10-May-2001 #text\_change 16-Aug-2004  
 C:Accession: A89813  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu-  
 ma, A.; Mizutani, O.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: A89813  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1499 <KUR>  
 A:Cross-references: UNIPROT:Q99WD1; GB:BA000018; PID:g13700362; PIDN:BA841660.1; GSPDB:  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: gltB  
 C:Superfamily: Glutamate synthase, large subunit

Query Match 38.2%; Score 43.5; DB 2; Length 1499;  
 Best Local Similarity 64.3%; Pred. No. 3.1e+02;  
 Matches 9; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 5 DGPTLRWISFCGG 18  
 Db 339 DGPTLRWISFCGG 349

```
RESULT 17
D72595
hypothetical protein APE1229 - Aeropyrum pernix (strain K1)
C/Dates: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Species: Aeropyrum pernix
C/Accession: D72595
R/Kawababyasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Hatake, Y., Jin-no, K., Takai-
awa, H., Takamiya, M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., K
DNA Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A/Reference number: A72450; MUID:9310339; PMID:10382966
A/Accession: D72595
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <KMW>
A/Cross-references: UNIPROT:Q9YCM9; DDBJ:AP000061; NID:95104821; PIDN:BAAB0218.1; PID:Q1
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE1229

Query Match          37.7%; Score 43; DB 2; Length 113;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 GPTLRWISFCGG 18
    |||:|||||
Db 21 GEARLRCWPSFCRG 33

RESULT 18
T15386
hypothetical protein CO3B1.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Dates: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T15386
R/Martin, U.
Submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid CO3B1.
A/Reference number: Z18340
A/Accession: T15386
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-115 <KAR>
A/Cross-references: UNIPROT:Q11110; EMBL:U40952; NID:91072237; PID:91072244; PIDN:AAA817
C/Genetics:
A/Gene: CESP:CO3B1.3
A/Introns: 80/1

Query Match          37.7%; Score 43; DB 2; Length 115;
Best Local Similarity 46.7%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCG 17
    |||:|||||
Db 68 CAGSEVHYHWACFCG 82

RESULT 19
A24902
erythropoietin precursor - mouse
C/Species: Mus musculus (house mouse)
C/Dates: 25-Oct-1997 #sequence_revision 15-Nov-1996 #text_change 09-Jul-2004
C/Accession: A24902; A24901
R/Shoemaker, C.B.; Miltsock, L.D.
Mol. Cell. Biol. 6, 849-858, 1986
A/Title: Murine erythropoietin gene: cloning, expression, and human gene homology.
A/Reference number: A24902; MUID:87039105; PMID:3773894
A/Accession: A24902
A/Molecule type: DNA
A/Residues: 1-192 <SHO>
A/Cross-references: UNIPROT:P07321
A/Note: the authors translated the codon TTA for residue 12 as Phe, TTA for residue 43 as
R; McDonald, J.D.; Lin, F.K.; Goldwasser, E.
```

```
Mol. Cell. Biol. 6, 842-848, 1986
A/Title: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin gene
A/Reference number: A24901; MUID:87039104; PMID:3022133
A/Accession: A24901
A/Molecule type: DNA
A/Residues: 1-67; 'P', 69-192 <MCD>
A/Cross-references: GB:M12930; NID:9193086; PIDN:AAA7570.1; PID:9387152
C/Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver c
C/Genetics:
A/Introns: 5/1; 52/3; 81/3; 141/3
C/Function:
A/Description: the primary inducer of erythrocyte formation
C/Superfamily: erythropoietin
C/Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F/1-26/Domain: signal sequence #status predicted <SIG>
F/27-192/Product: erythropoietin #status predicted <MAT>
F/33-187,55-165/Disulfide bonds: #status predicted
F/50,64,109/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match          37.7%; Score 43; DB 1; Length 192;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCADGPTLRWISFCGK 19
    |||:|||||
Db 54 GCAEPRLENTVADTK 71

RESULT 20
I46885
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Dates: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I46885; S43172
R/Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A/Title: Genes for mast-cell serine protease and their molecular evolution.
A/Reference number: I46884; MUID:95048582; PMID:7959952
A/Accession: I46885
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-230 <RES>
A/Cross-references: UNIPROT:P1845; EMBL:X78542; NID:9468809; PIDN:CAA55288.1; PID:9468
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase
F/32-230/Domain: trypsin homology #status atypical <TRY>

Query Match          37.7%; Score 43; DB 2; Length 230;
Best Local Similarity 70.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 LRWISFCGG 18
    |||:|||||
Db 53 LNYWIHFCCG 62

RESULT 21
AB2768
lipote liposynthesis protein B [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C/Species: Agrobacterium tumefaciens
C/Dates: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AB2768
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guentherer, D.; Kulyavin, T.; Levy, R.; Li, M.; Mclel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AB2768
A/Status: preliminary
A/Molecule type: DNA
```



A;Residues: 1-233 <KIR>  
A;Cross-references: UNIPROT:Q8UF44; GB:AE008688; PIDN:AA42560.1; PID:917739983; GSPDB:G  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: 11pB  
A;Map position: circular chromosome  
C;Superfamily: Becherichia coli lipocate-protein ligase 11pB

Query Match 37.7%; Score 43; DB 2; Length 233;  
Best Local Similarity 43.5%; Pred. No. 66;  
Matches 10; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 GGCAD-----GPTREWISFCG 17  
DB 148 GGMADKIALGIRLRKWSFPG 170

RESULT 22  
T19988  
hypothetical protein C47B2.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19988  
R;Kershaw, J.  
submitted to the EMBL Data Library, October 1997  
A;Reference number: Z19208  
A;Accession: T19988  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-246 <WIL>  
A;Cross-references: UNIPROT:O62106; EMBL:Z99709; PIDN:CA816860.1; GSPDB:GN00019; CESP:CA  
A;Experimental source: clone C47B2  
C;Genetics:  
A;Gene: CESP:C47B2.5  
A;Map position: 1  
A;Intons: 91/3; 127/3  
C;Superfamily: conserved hypothetical protein YPR016c

Query Match 37.7%; Score 43; DB 2; Length 246;  
Best Local Similarity 41.7%; Pred. No. 70;  
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 GPTREWISFCG 17  
DB 196 GMYVNDWVAFPG 207

RESULT 23  
T01012  
probable translation initiation factor [imported] - Arabidopsis thaliana  
N;Alternate names: hypothetical protein T517.12  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: T01012; H84821  
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, November 1997  
A;Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.  
A;Reference number: Z14152  
A;Accession: T01012  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-247 <ROU>  
A;Cross-references: UNIPROT:O22290; EMBL:AC003000; NID:g2642152; PIDN:AA87131.1; PID:g2  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Koffel, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, J.; Taiton, L.  
euser, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: H84821  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-247 <SNO>  
A;Cross-references: GB:AE002093; NID:g2642164; PIDN:AA87131.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: T517.12; At2g39820  
A;Map position: 2  
A;Intons: 4/1; 38/2; 82/1; 162/3  
C;Superfamily: conserved hypothetical protein YPR016c

Query Match 37.7%; Score 43; DB 2; Length 247;  
Best Local Similarity 50.0%; Pred. No. 70;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 GPTREWISFCG 17  
DB 198 GLTVNDWTAFPG 209

RESULT 24  
D97548  
lipocate-protein ligase b (lipocate biosynthesis protein b) [imported] - Agrobacterium tum  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: D97548  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: D97548  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-268 <KIR>  
A;Cross-references: UNIPROT:Q8UF44; GB:AE007869; PIDN:AA87341.1; PID:915156641; GSPDB:G  
A;Genetics:  
A;Gene: AGR\_C\_2865  
A;Map position: circular chromosome  
C;Superfamily: Becherichia coli lipocate-protein ligase 11pB

Query Match 37.7%; Score 43; DB 2; Length 268;  
Best Local Similarity 43.5%; Pred. No. 75;  
Matches 10; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 1 GGCAD-----GPTREWISFCG 17  
DB 183 GGMADKIALGIRLRKWSFPG 205

RESULT 25  
A38654  
mast cell proteinase 6 (BC 3.4.21.-) precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 21-Feb-1992 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: A38654; B38654; B3646; I59478  
R;Remoldes, D.S.; Gurley, D.S.; Austen, K.F.; Serafin, W.E.  
J. Biol. Chem. 266, 3847-3853, 1991  
A;Title: Cloning of the cDNA and gene of mouse mast cell protease-6. Transcription by p  
A;Reference number: A38654; MUID:91139682; PMID:1995638  
A;Accession: A38654  
A;Molecule type: DNA  
A;Residues: 1-276 <REV>  
A;Cross-references: UNIPROT:P21845; GB:M57625; NID:g200506; PIDN:AA39987.1; PID:g20050  
A;Note: the authors translated the codon GGC for residue 24 as Ala, GAG for residue 37  
as Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly  
A;Accession: B38654  
A;Molecule type: mRNA  
A;Residues: 1-276 <RE2>  
A;Cross-references: GB:M57626; NID:g200508; PIDN:AA39988.1; PID:g200509  
R;Remoldes, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990  
A;Title: Different mouse mast cell populations express various combinations of at least  
A;Reference number: A35646; MUID:90222202; PMID:2326280  
A;Accession: D35646  
A;Status: preliminary  
A;Molecule type: protein

A;Residues: 32-54 <RES>  
R;Huang, R.; Abirink, M.; Gohl, A.E.; Nilsson, G.; Aveekogoh, M.; Larsson, L.G.; Nilsson, Scand. J. Immunol. 38, 359-367, 1993  
A;Title: Expression of a mast cell tryptase in the human monocytic cell lines U-937 and  
A;Reference number: 159478; MUID:94023807; PMID:8210998  
A;Accession: 159478  
A;Status: preliminary; translated from GB/EMBL/DBD  
A;Molecule type: mRNA  
A;Residues: 1-276 <RES>  
A;Cross-references: GB:U31853; NID:9473480; PIDN:AAA39725.1; PID:9473481  
C;Genetics:  
A;Gene: MWC6-6  
A;Intons: 24/1; 79/2; 168/1; 222/3  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase; zymogen  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-33/Domain: activation peptide #status predicted <ACT>  
F;33-276/Product: mast cell proteinase 6 #status experimental <MAT>  
F;33-268/Domain: trypsin homology <TRY>  
F;75,122,225/Active site: His, Asp, Ser #status predicted

Query Match 37.7%; Score 43; DB 2; Length 276;  
Best Local Similarity 70.0%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 9 LREWISFCGK 18  
| | | | |  
Db 53 LNWVHFCG 62

RESULT 26  
S21324  
Probable beta-glucosidase - Ruminococcus flavefaciens  
C;Species: Ruminococcus flavefaciens  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S21324  
R;Huang, C.M.; Amundson, R.V.; Yu, P.L.  
submitted to the EMBL Data Library, September 1990  
A;Description: Nucleotide sequence of a cellulase gene complex from Ruminococcus flavefa  
A;Reference number: S21323  
A;Accession: S21324  
A;Molecule type: DNA  
A;Residues: 1-434 <HUA>  
A;Cross-references: UNIPROT:Q52748; EMBL:X56082; NID:946152; PIDN:CAA39560.1; PID:e33392  
A;Note: the authors designated this protein as beta-glucosidase

Query Match 37.7%; Score 43; DB 2; Length 434;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 6 GPTLEWISFCGK 19  
| | | | |  
Db 75 GPSYGYWTNCGK 88

RESULT 27  
AE2497  
hypothetical protein alr7157 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AE2497  
R;Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasaoto, S.; Watanabe, A.; Iriyuchi, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Accession number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AE2497  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-629 <KUR>  
A;Cross-references: UNIPROT:O8YK71; GB:BA000020; PIDN:BA078241.1; PID:g17135695; GSPDB:C

A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr7157  
A;Genome: plasmid

Query Match 37.7%; Score 43; DB 2; Length 629;  
Best Local Similarity 54.5%; Pred. No. 1.7e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 9 LREWISFCGK 19  
| | | | |  
Db 215 LKQWQNFCAKG 225

RESULT 28  
S54478  
probable membrane protein YMR266w - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YMR156.08  
C;Species: Saccharomyces cerevisiae  
C;Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C;Accession: S54478  
R;Lye, G.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S54014  
A;Accession: S54478  
A;Molecule type: DNA  
A;Residues: 1-953 <LYE>  
A;Cross-references: UNIPROT:Q03516; EMBL:249260; NID:9809081; PID:9809089; GSPDB:GN0001.  
A;Experimental source: strain AB972  
C;Genetics:  
A;Gene: SGD:RSN1; MIPS:YMR266w  
A;Cross-references: SGD:S0004879  
A;Map position: 13R  
C;Superfamily: Yeast probable membrane protein YOL084w  
C;Keywords: transmembrane protein  
F;33-48/Domain: transmembrane #status predicted <TM1>  
F;106-122/Domain: transmembrane #status predicted <TM2>  
F;152-168/Domain: transmembrane #status predicted <TM3>  
F;395-411/Domain: transmembrane #status predicted <TM4>  
F;435-451/Domain: transmembrane #status predicted <TM5>  
F;545-561/Domain: transmembrane #status predicted <TM6>  
F;599-615/Domain: transmembrane #status predicted <TM7>  
F;646-662/Domain: transmembrane #status predicted <TM8>  
F;668-684/Domain: transmembrane #status predicted <TM9>

Query Match 37.7%; Score 43; DB 2; Length 953;  
Best Local Similarity 41.2%; Pred. No. 2.4e+02;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 GGCADGPTLEWISFCG 17  
| | | | |  
Db 558 GAFIDGTRKKNRFG 574

RESULT 29  
B37227  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 09-Jul-2004  
C;Accession: B37227; I50395  
R;Takeyasu, K.; Lemae, V.; Fambrough, D.M.  
Am. J. Physiol. 259, C619-C630, 1990  
A;Title: Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.  
A;Reference number: A37227; MUID:91023019; PMID:2171348  
A;Accession: B37227  
A;Molecule type: mRNA  
A;Residues: 1-1010 <TA2>  
A;Cross-references: UNIPROT:P24798; GB:M59960; NID:g212407; PIDN:AAA48982.1; PID:g21240  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; sodium t  
F;574-770/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;202-470/Binding site: carbonylate (Asn) (covalent) #status predicted  
F;363/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:495/Binding site: ATP (Lys) #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1010;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 7 PTLREWSFC 16  
Db 71 PTPREWKFC 80

# RESULT 30

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - human

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text\_change 09-Jul-2004

C:Accession: S00801; S04019; A27397; S02275

R:Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Melkov, A.M.; S

dyanov, N.N.; Sverdlov, E.D.

FEBS Lett. 233, 87-94, 1988

A:Title: Family of human Na,K-ATPase genes. Structure of the gene for the catalytic sub

A:Reference number: S00801; MUID:88255304; PMID:2838329

A:Accession: S00801

A:Molecule type: DNA

A:Residues: 1-1013 <OVC>

A:Cross-references: UNIPROT:P13637; EMBL:M37456

R:Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Melkov, A.M.; Smir

ov, N.N.; Ovchinnikov, Y.A.

Dokl. Biochem. 297, 426-431, 1987

A:Title: Family of human Na(+),K(+)-ATPase genes. Structure of the gene of isoform alpha

A:Reference number: S04019

A:Accession: S04019

A:Molecule type: DNA

A:Residues: 1, 'EIH', '3-1013 <SVEI>

A:Cross-references: EMBL:X1910; NID:928963

A>Note: The authors translated the codon TTC for residue 283 as Ser and TCT for residue

A>Note: this paper is a translation of the Russian paper published in Dokl. Akad. Nauk S

R:Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Allikmets, R.L.; M

lina, M.B.; Sverdlov, V.E.; Modyanov, N.N.; Ovchinnikov, Y.A.

FEBS Lett. 217, 275-278, 1987

A:Title: The family of human Na(+),K(+)-ATPase genes. No less than five genes and/or pseudog

A:Reference number: A27397; MUID:87247232; PMID:3035682

A:Accession: A27397

A:Molecule type: mRNA

A:Residues: 243-434 <SVE2>

A:Cross-references: GB:M27570

C:Genetics:

A:Gene: GDB:ATPLA3

A:Cross-references: GDB:119713; OMIM:182350

A:Map position: 19q13.2-19q13.2

A:Introns: 2/3; 31/3; 51/3; 119/3; 157/3; 202/3; 242/1; 331/3; 398/1; 435/2; 479/3; 544/

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

F:86-110/Domain: transmembrane #status predicted <TM1>

F:120-139/Domain: transmembrane #status predicted <TM2>

F:140-260/Domain: intracellular #status predicted <INT2>

F:281-303/Domain: transmembrane #status predicted <TM3>

F:330-338/Domain: transmembrane #status predicted <TM4>

F:339-776/Domain: intracellular #status predicted <INT3>

F:577-773/Domain: ATPase nucleotide-binding domain homology <ATN>

F:777-800/Domain: transmembrane #status predicted <TM5>

F:839-864/Domain: transmembrane #status predicted <TM6>

F:965-942/Domain: intracellular #status predicted <INT4>

F:943-968/Domain: transmembrane #status predicted <TM7>

F:965-1013/Domain: extracellular #status predicted <EXT>

F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:498/Binding site: Asp (Lys) #status predicted

F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 1; Length 1013;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 7 PTLREWSFC 16  
Db 74 PTPREWKFC 83

# RESULT 31

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - rat

N:Alternate names: Na+/K+-transporting ATPase alpha (III) chain

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Jun-1988 #sequence revision 23-Apr-1993 #text\_change 09-Jul-2004

C:Accession: C24639; S00514; B27180; A60470

R:Shull, G.E.; Greeb, J.; Lingrel, J.B.

Biochemistry 25, 8125-8132, 1986

A:Title: Molecular cloning of three distinct forms of the Na(+),K(+)-ATPase alpha-subunit f

A:Reference number: A90512; MUID:87128908; PMID:3028470

A:Accession: C24639

A:Molecule type: mRNA

A:Residues: 1-1013 <SHU>

A:Cross-references: UNIPROT:P06687; EMBL:M14513; NID:9203030; PIDN:AAA40777.1; PID:9203

A>Note: In the authors' translation 405-Ser is shown after residue 409 and, consequentl

R:Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.;

J. Biochem. 102, 43-58, 1987

A:Title: Primary structures of two types of alpha-subunit of rat brain Na(+),K(+)-ATPase

A:Reference number: S00460; MUID:88032933; PMID:2822682

A:Accession: S00514

A:Molecule type: mRNA

A:Residues: 1-907, 'C', '909-1013 <HAR>

A:Cross-references: EMBL:X05883; NID:955769; PIDN:CAA29307.1; PID:955770

R:Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.

J. Cell Biol. 105, 1855-1865, 1987

A:Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural

A:Reference number: A92749; MUID:88033255; PMID:2822726

A:Accession: B27180

A:Molecule type: mRNA

A:Residues: 1, 'NL', '4-103', 'R', '105-113', 'E', '115-127', 'G', '129-148', 'O', '150-151', 'T', '153-165', 'D

A:Cross-references: EMBL:M26648; NID:9205633; PIDN:AAA1672.1; PID:9205634

A>Note: The authors translated the codon CAG for residue 149 as Glu, GGC for residue 19

R:Han, Y.M.; Goldict, G.

Biochemistry 28, 569-573, 1989

A:Title: Rat brain has the alpha3 form of the (Na(+),K(+)-ATPase.

A:Reference number: A60470; MUID:89229049; PMID:2540801

A:Accession: A60470

A:Molecule type: protein

A:Residues: 117-132;586-595, 'X', '597-601 <HSU>

C:Comment: The alpha-3 form appears to be highly ouabain-inhibitable, as is alpha-2 but

C:Genetics:

A:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans

F:86-110/Domain: transmembrane #status predicted <TM1>

F:120-139/Domain: transmembrane #status predicted <TM2>

F:140-280/Domain: intracellular #status predicted <INT2>

F:281-303/Domain: transmembrane #status predicted <TM3>

F:330-338/Domain: transmembrane #status predicted <TM4>

F:339-776/Domain: intracellular #status predicted <INT3>

F:577-773/Domain: ATPase nucleotide-binding domain homology <ATN>

F:777-800/Domain: transmembrane #status predicted <TM5>

F:839-864/Domain: transmembrane #status predicted <TM6>

F:965-942/Domain: intracellular #status predicted <INT4>

F:943-968/Domain: transmembrane #status predicted <TM7>

F:965-1013/Domain: extracellular #status predicted <EXT>

F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:498/Binding site: Asp (Lys) #status predicted

F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1013;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 7 PTLREWSFC 16  
Db 74 PTPREWKFC 83

## RESULT 32

A37227  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 16-Sep-1992 #sequence revision 13-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: I50394; A37227  
R/KeyAsn, K.; Lemae, M.; Fambrough, D.M.  
Am. J. Physiol. 259, 619-630, 1991  
A/Title: Stability of the Na+, K+-ATPase alpha-subunit isoforms in evolution.  
A/Reference number: I50394  
A/Accession: I50394  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1017 <TA>  
A/Cross-references: UNIPROT:P24797; GB:M59599; NID:g212405; PIDN:AAA48981.1; PID:g212406  
R/KeyAsn, K.; Lemae, V.; Fambrough, D.M.  
Am. J. Physiol. 259, C619-C630, 1990  
A/Title: Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.  
A/Reference number: A37227; MUID:91023019; PMID:2171348  
A/Accession: A37227  
A/Molecule type: mRNA  
A/Residues: 3-1017 <TA>  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; glycoprotein; hydrolase; phosphoprotein  
F:581-777/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:210-478/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:371/Active site: Asp (aspartylphosphate intermediate) #status predicted

## Query Match

Best Local Similarity 37.7%; Score 43; DB 2; Length 1017;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWISFC 16  
DB 79 PTLREWVKFC 88

## RESULT 33

A34474  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - human  
N/Alternate names: Na+/K+-exchanging ATPase alpha chain-4; sodium/potassium transporting  
C/Species: Homo sapiens (man)  
C/Date: 15-Jun-1990 #sequence revision 15-Jun-1990 #text\_change 09-Jul-2004  
C/Accession: A34474; B27795; D27397  
R/Shull, M.M.; Pugh, D.G.; Lingrel, J.B.  
J. Biol. Chem. 264, 17532-17543, 1989  
A/Title: Characterization of the human Na, K-ATPase alpha2 gene and identification of int  
A/Reference number: A34474; MUID:9008924; PMID:2477373  
A/Accession: A34474  
A/Molecule type: DNA  
A/Residues: 1-1020 <SHU>  
A/Cross-references: UNIPROT:P50993; GB:J05096; NID:g179164; PIDN:AAA51797.1; PID:g179165  
R/Shull, M.M.; Lingrel, J.B.  
Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
A/Title: Multiple genes encode the human Na+, K+-ATPase catalytic subunit.  
A/Reference number: A94158; MUID:87231946; PMID:3035563  
A/Accession: B27795  
A/Molecule type: DNA  
A/Residues: 211-249 <SH2>  
A/Cross-references: GB:M16795; NID:g179196; PIDN:AAA51799.1; PID:g553194  
R/Sverdlov, E.D.; Monastyrskaya, G.S.; Brode, N.E.; Ushkaryov, Y.A.; Altkmetz, R.L.; M  
lina, M.B.; Sverdlov, V.E.; Modyanov, N.N.; Ovchinnikov, Y.A.  
FEBS Lett. 217, 275-278, 1987  
A/Title: The family of human Na+, K+-ATPase genes. No less than five genes and/or pseudog  
A/Reference number: A27397; MUID:8724723; PMID:3036582  
A/Accession: D27397  
A/Molecule type: DNA  
A/Residues: 251-442 <SVE>  
A/Cross-references: GB:M27571  
C/Genetics:  
A/Gene: GDB:ATP1A2

A/Cross-references: GDB:119712; OMIM:182340

A/Map position: 1q21-1q23  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
F:6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MNT>  
F:6-93/Domain: intracellular #status predicted <INT1>  
F:94-118/Domain: transmembrane #status predicted <TM1>  
F:128-147/Domain: transmembrane #status predicted <TM2>  
F:148-288/Domain: intracellular #status predicted <INT2>  
F:289-311/Domain: transmembrane #status predicted <TM3>  
F:318-346/Domain: transmembrane #status predicted <TM4>  
F:347-783/Domain: intracellular #status predicted <INT3>  
F:584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:784-807/Domain: transmembrane #status predicted <TM5>  
F:846-871/Domain: transmembrane #status predicted <TM6>  
F:872-949/Domain: intracellular #status predicted <INT4>  
F:950-975/Domain: transmembrane #status predicted <TM7>  
F:976-1020/Domain: extracellular #status predicted <EXT>  
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:505/Binding site: Asp (Lys) #status predicted  
F:714,718,723/Active site: Asp, Asp, Lys #status predicted

## Query Match

Best Local Similarity 37.7%; Score 43; DB 2; Length 1020;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWISFC 16  
DB 82 PTLREWVKFC 91

## RESULT 34

B24639  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - rat  
N/Alternate names: Na+/K+-transporting ATPase alpha-plus chain  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text\_change 09-Jul-2004  
C/Accession: B24639  
R/Shull, G.E.; Greb, J.; Lingrel, J.B.  
Biochemistry 25, 8125-8132, 1986  
A/Title: Molecular cloning of three distinct forms of the Na+, K+-ATPase alpha-subunit f  
A/Reference number: A90512; MUID:87128908; PMID:3028470  
A/Accession: B24639  
A/Molecule type: mRNA  
A/Residues: 1-1020 <SHU>  
A/Cross-references: UNIPROT:P06686; EMBL:M14512; NID:g203028; PIDN:AAA40776.1; PID:g203  
C/Genetics:  
A/Gene: NIKRA2  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans  
F:6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MNT>  
F:6-93/Domain: intracellular #status predicted <INT1>  
F:94-119/Domain: transmembrane #status predicted <TM1>  
F:128-147/Domain: transmembrane #status predicted <TM2>  
F:148-288/Domain: intracellular #status predicted <INT2>  
F:289-311/Domain: transmembrane #status predicted <TM3>  
F:318-346/Domain: transmembrane #status predicted <TM4>  
F:347-783/Domain: intracellular #status predicted <INT3>  
F:584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:784-807/Domain: transmembrane #status predicted <TM5>  
F:846-871/Domain: transmembrane #status predicted <TM6>  
F:872-949/Domain: intracellular #status predicted <INT4>  
F:950-975/Domain: transmembrane #status predicted <TM7>  
F:976-1020/Domain: extracellular #status predicted <EXT>  
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:505/Binding site: Asp (Lys) #status predicted  
F:714,718,723/Active site: Asp, Asp, Lys #status predicted

## Query Match

Best Local Similarity 37.7%; Score 43; DB 2; Length 1020;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 PTLREWISFC 16

Db 82 PTPBWKFC.91

## RESULT 35

PMSHNA

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain precursor - sheep  
N/Alternate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha chain  
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C/Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004

C/Accession: A01074; A35426  
R/Shul, G.E.; Schwartz, A.; Lingrel, J.B.  
Nature 316: 691-695, 1985

A/Title: Amino-acid sequence of the catalytic subunit of the (Na(+)-K(+)) ATPase deduced  
A/Reference number: A01074; MUID:85296299; PMID:2993903

A/Accession: A01074  
A/Molecule type: mRNA  
A/Residues: 1-1021 <SHU>

A/Cross-references: UNIPROT:P04074; GB:X02813; NID:g1205; PIDN:CAA26581.1; PID:g1206  
J. Biol. Chem. 265, 10260-10265, 1990

A/Title: Lysine 480 is an essential residue in the putative ATP site of lamb kidney (Na,  
A/Reference number: A35426; MUID:90285144; PMID:2162343

A/Accession: A35426  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 475-492 <HIN>

A/Comment: This is the catalytic component of the active enzyme, which catalyzes the hydrolysis of ATP to ADP and inorganic phosphate, providing the energy for active transport of sodium and potassium across the cell membrane.

C/Comment: This enzyme is specifically inhibited by cardiac glycosides such as digoxin and ouabain.  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; hydrolyase; phosphoprotein; potassium transport; sodium transport; trans

F:6-1021/Product: Na+/K+-transporting ATPase alpha chain #status predicted <MAT>  
F:94-115/Domain: transmembrane #status predicted <TM1>  
F:128-144/Domain: transmembrane #status predicted <TM2>

F:289-311/Domain: transmembrane #status predicted <TM3>  
F:318-346/Domain: transmembrane #status predicted <TM4>  
F:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>

F:785-808/Domain: transmembrane #status predicted <TM5>  
F:847-872/Domain: transmembrane #status predicted <TM6>  
F:951-976/Domain: transmembrane #status predicted <TM7>

F:315/Binding site: cardiac glycoside (TTP) #status predicted  
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:506/Binding site: ATP (Lys) #status predicted

Query Match 37.7%; Score 43; DB 1; Length 1021;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 82 PTPBWKFC 91

## RESULT 36

S04630

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - horse

C/Species: Equus caballus (domestic horse)  
C/Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004

C/Accession: S04630  
R/Kano, I.; Nagai, F.; Satoh, K.; Ushiyama, K.; Nakao, T.; Kano, K.  
FEBS Lett. 250, 91-98, 1989

A/Title: Structure of the alpha(1) subunit of horse Na,K-ATPase gene.  
A/Reference number: S04630; MUID:89290042; PMID:2544461

A/Accession: S04630  
A/Molecule type: DNA  
A/Residues: 1-1021 <KAN>

A/Cross-references: UNIPROT:P18907; EMBL:X16773; NID:g1010; PIDN:CAA34716.1; PID:9871026  
C/Genetics: 4/3; 39/3; 59/3; 127/3; 165/3; 210/3; 250/1; 339/3; 406/1; 442/3; 487/3; 552/3  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

F:6-1021/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>

F:94-118/Domain: transmembrane #status predicted <TM1>  
F:128-147/Domain: transmembrane #status predicted <TM2>

F:289-311/Domain: transmembrane #status predicted <TM3>  
F:318-346/Domain: transmembrane #status predicted <TM4>  
F:347-374/Domain: transmembrane #status predicted <TM5>

F:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:785-808/Domain: transmembrane #status predicted <TM6>  
F:847-872/Domain: transmembrane #status predicted <TM7>

F:873-950/Domain: intracellular #status predicted <INT4>  
F:951-976/Domain: transmembrane #status predicted <TM7>  
F:977-1021/Domain: extracellular #status predicted <EXT>

F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:506/Binding site: ATP (Lys) #status predicted  
F:715,719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 1; Length 1021;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 82 PTPBWKFC 91

## RESULT 37

A28199

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - chicken

C/Species: Gallus gallus (chicken)  
C/Date: 21-Sep-1988 #sequence\_revision 21-Sep-1988 #text\_change 09-Jul-2004

C/Accession: A28199  
R/Takeyasu, K.; Tamkun, M.W.; Renaud, K.J.; Fambrough, D.M.  
J. Biol. Chem. 263, 4347-4354, 1988

A/Title: Ouabain-sensitive (Na(+)+K(+))-ATPase activity expressed in mouse L cells by  
A/Reference number: A28199; MUID:88153759; PMID:2831227

A/Accession: A28199  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA

A/Cross-references: UNIPROT:P09572; GB:J03230; NID:g211219; PIDN:AAA48607.1; PID:g21122  
C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C/Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein

F:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:213,481/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:506/Binding site: ATP (Lys) #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1021;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 82 PTPBWKFC 91

## RESULT 38

B24862

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - pig

N/Alternate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha chain  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004

C/Accession: B24862; A35504; S00011; S00502; S02569; S29762  
R/Ovchinnikov, Y.A.; Modayany, N.N.; Broude, N.E.; Petrunkin, K.B.; Arzai  
FEBS Lett. 201, 237-245, 1986

A/Title: Pig kidney Na+,K+-ATPase. Primary structure and spatial organization.  
A/Reference number: A91361; MUID:86220813; PMID:2423371

A/Accession: B24862  
A/Molecule type: mRNA  
A/Residues: 1-1021 <OVCH2>  
A/Cross-references: UNIPROT:P05024; EMBL:X03938; NID:g1897; PIDN:CAA27576.1; PID:g1898

A/Note: the authors translated the codon TCC for residue 391 as Phe. TCG for residue 723  
 A/Note: part of this sequence, including the amino and carboxyl end of the mature protein  
 R/Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Arsenyan, S.G.; Brode, N.E.; Petrkhin, K.E.;  
 Dohl. Biochem. 283, 270-272, 1985  
 A/Title: Amino acid sequence of the 17-kilodalton fragment of the cytoplasmic region of  
 A/Reference number: 146572  
 A/Accession: 146572  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 469-617 <OVCL>  
 A/Cross-references: GB:M32512; NID:g164385; PIDN:AAA1004.1; PID:g164386  
 R/Karlish, S.J.D.; Goldshleger, R.; Stein, W.D.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 4566-4570, 1990  
 A/Title: A 19-kDa C-terminal tryptic fragment of the alpha chain of Na/K-ATPase is essen  
 A/Reference number: A35504; MUID:90280416; PMID:2162048  
 A/Accession: A35504  
 A/Molecule type: protein  
 A/Residues: 836-845, 'R', 847-851 <KAR>  
 R/Ovchinnikov, Y.A.; Arzamazova, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Aldanova, N.  
 FEBS Lett. 217, 269-274, 1987  
 A/Title: Detailed structural analysis of exposed domains of membrane-bound Na<sup>+</sup>, K<sup>+</sup>-ATPase  
 A/Reference number: S00011; MUID:87247231; PMID:3036581  
 A/Contents: annotation; membrane topology  
 R/Ovchinnikov, Y.A.; Luneva, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Arzamazova, N.M.  
 FEBS Lett. 227, 230-234, 1988  
 A/Title: Topology of Na<sup>+</sup>, K<sup>+</sup>-ATPase: identification of the extra- and intracellular hydrop  
 A/Reference number: S02569; MUID:88112252; PMID:2448169  
 A/Contents: annotation; membrane topology  
 C/Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C/Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F/6-1021/Product: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain #status experimental <MAT>  
 F/6-93/Domain: intracellular #status predicted <INT1>  
 F/94-118/Domain: transmembrane #status predicted <TM1>  
 F/128-147/Domain: transmembrane #status predicted <INT2>  
 F/168-268/Domain: intracellular #status predicted <INT2>  
 F/269-311/Domain: transmembrane #status predicted <TM>  
 F/318-346/Domain: transmembrane #status predicted <TM>  
 F/347-784/Domain: intracellular #status predicted <INT3>  
 F/585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F/785-808/Domain: transmembrane #status predicted <TM5>  
 F/847-872/Domain: transmembrane #status predicted <TM6>  
 F/951-950/Domain: intracellular #status predicted <INT4>  
 F/977-1021/Domain: extracellular #status predicted <EXT>  
 F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F/505/Binding site: ATP (Lys) #status predicted  
 F/716,719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1021;  
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 7 PTLREWISFC 16  
 DB 82 PTLREWVKFC 91

RESULT 39  
 Na<sup>+</sup>/K<sup>+</sup>-exchanging ATPase (EC 3.6.3.9) alpha chain - European eel  
 C/Species: Anguilla anguilla (European eel)  
 C/Date: 01-Feb-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004  
 C/Accession: S49127  
 R/Cutler, C.; Sanders, J.L.; Cramb, G.  
 submitted to the EMBL Data Library, November 1993  
 A/Reference number: S45093  
 A/Accession: S49127  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1022 <CUT>  
 A/Cross-references: UNIPROT:Q92030; EMBL:X76108; NID:g509405; PIDN:CAA53714.1; PID:g5094  
 C/Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C/Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; transmem

F/586-782/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F/214,482/Binding site: carboxylate (asn) (covalent) #status predicted  
 F/375/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F/507/Binding site: ATP (Lys) #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1022;  
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 7 PTLREWISFC 16  
 DB 83 PTLREWVKFC 92

RESULT 40  
 Na<sup>+</sup>/K<sup>+</sup>-exchanging ATPase (EC 3.6.3.9) alpha-1 chain [validated] - rat  
 N/Alternate names: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain, kidney-type  
 N/Contents: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha-s chain  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
 C/Accession: A24639; S00460; A27180; S11020; A25171; S29877; S10758  
 R/Shull, G.E.; Greb, J.; Lingrel, J.B.  
 Biochemistry 25, 8125-8132, 1986  
 A/Title: Molecular cloning of three distinct forms of the Na<sup>+</sup>, K<sup>+</sup>-ATPase alpha-subunit f.  
 A/Reference number: A90512; MUID:87128908; PMID:3028470  
 A/Accession: A24639  
 A/Molecule type: mRNA  
 A/Residues: 1-1023 <SHU>  
 A/Cross-references: UNIPROT:P06685; EMBL:M14511; NID:g203026; PIDN:AAA40775.1; PID:g2030  
 R/Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.; f  
 J. Biochem. 102, 43-58, 1987  
 A/Title: Primary structures of two types of alpha-subunit of rat brain Na(+) ,K(+) -ATPase  
 A/Reference number: S00460; MUID:88032933; PMID:2822682  
 A/Accession: S00460  
 A/Molecule type: mRNA  
 A/Residues: 1-1023 <HAR>  
 A/Cross-references: EMBL:X05882; NID:g55771; PIDN:CAA29306.1; PID:g55772  
 R/Heriera, V.U.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.  
 J. Cell Biol. 105, 1855-1865, 1987  
 A/Title: Three differentially expressed Na<sup>+</sup>, K<sup>+</sup>-ATPase alpha subunit isoforms: structural  
 A/Reference number: A92749; MUID:88033255; PMID:2822726  
 A/Accession: A27180  
 A/Molecule type: mRNA  
 A/Residues: 1-67, 'PV', 70-174, 'E', 176-187, 'V', 189-334, 'V', 336-1023 <HER>  
 A/Reference number: 1-67, 'PV', 70-174, 'E', 176-187, 'V', 189-334, 'V', 336-1023 <HER>  
 A/Cross-references: EMBL:M28647; NID:g205631; PIDN:AAA41671.1; PID:g205632  
 R/Yagawa, Y.; Kawakami, K.; Negano, K.  
 Biochim. Biophys. Acta 1049, 286-292, 1990  
 A/Title: Cloning and analysis of the 5'-flanking region of rat Na(+) /K(+) -ATPase alpha-  
 A/Reference number: S11020; MUID:90344872; PMID:2166579  
 A/Accession: S11020  
 A/Status: translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-41 <YAG>  
 A/Cross-references: EMBL:X53233  
 R/Schneider, J.W.; Mercer, R.W.; Caplan, M.; Emanuel, J.R.; Sweadner, K.J.; Benz Jr., B.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6357-6361, 1985  
 A/Title: Molecular cloning of rat brain Na<sup>+</sup>, K<sup>+</sup>-ATPase alpha-subunit cDNA.  
 A/Reference number: A25171; MUID:85298352; PMID:2994074  
 A/Accession: A25171  
 A/Molecule type: mRNA  
 A/Residues: 489-533 <SCH>  
 R/Lytton, J.  
 Biochem. Biophys. Res. Commun. 132, 764-769, 1985  
 A/Title: The catalytic subunits of the Na(+) ,K(+) -ATPase alpha and alpha(+) isozymes  
 A/Reference number: S29877; MUID:8605067; PMID:2998384  
 A/Accession: S29877  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 6-19 <LYT>  
 R/Kurihara, K.; Hosoi, K.; Kodama, A.; Ueha, T.  
 Biochim. Biophys. Acta 1039, 234-240, 1990  
 A/Title: A new electrophoretic variant of alpha subunit of Na(+) /K(+) -ATPase from the B



A:Reference number: S10758; MUID:90304196; PMID:2163680  
A:Accession: S10758  
A:Molecule type: protein  
A:Residues: 6 'X', 8-10, 'X', 12-16 <KUR>  
A:Experimental source: submandibular gland  
A>Note: designated alpha-S form, thought to arise from alpha-1 chain by post-translational C:Genetics:  
A:Gene: NKAAL  
A:Introns: 4/3  
A>Note: the list of introns may be incomplete  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP, heterodimer, hydrolase, ion transport, phosphoprotein, potassium transp  
F:6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status experimental <MAT>  
F:96-120/Domain: intracellular #status predicted <INT1>  
F:96-120/Domain: transmembrane #status predicted <TM1>  
F:130-149/Domain: transmembrane #status predicted <TM2>  
F:150-280/Domain: intracellular #status predicted <INT2>  
F:291-313/Domain: transmembrane #status predicted <TM3>  
F:320-348/Domain: transmembrane #status predicted <TM4>  
F:349-786/Domain: intracellular #status predicted <INT3>  
F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:787-810/Domain: transmembrane #status predicted <TM5>  
F:849-874/Domain: transmembrane #status predicted <TM6>  
F:875-952/Domain: intracellular #status predicted <INT4>  
F:953-978/Domain: transmembrane #status predicted <TM7>  
F:979-1023/Domain: extracellular #status predicted <EXT>  
F:376/Active site: Asp (Aspartylphosphate intermediate) #status predicted  
F:508/Binding site: Asp (Lys) #status predicted  
F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 1; Length 1023;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 84 PTPREWVFC 93

RESULT 41  
S24650  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - giant toad  
C:Species: Bufo marinus (giant toad)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
C:Accession: A43451; S24650  
R:Jaisser, F.; Canessa, C.M.; Horibarger, J.D.; Rossier, B.C.  
J. Biol. Chem. 267, 16895-16903, 1992  
A:Title: Primary sequence and functional expression of a novel ouabain-resistant Na,K-AT  
A:Reference number: A43451; MUID:92380991; PMID:1380956  
A:Accession: A43451  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1023 <JAI>  
A:Cross-references: UNIPROT:P30714; EMBL:Z11798; NID:g62491; PIDN:CAA77642.1; PID:g62492  
A:Experimental source: urinary bladder cell line TEM 18-23  
A>Note: submitted to the EMBL Data Library, March 1992  
A>Note: sequence extracted from NCBI backbone (NCBIP:111876)  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP, heterodimer, hydrolase, ion transport, phosphoprotein, potassium transp  
F:6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>  
F:96-120/Domain: transmembrane #status predicted <INT1>  
F:96-120/Domain: transmembrane #status predicted <TM1>  
F:150-280/Domain: intracellular #status predicted <INT2>  
F:291-313/Domain: transmembrane #status predicted <TM2>  
F:320-348/Domain: transmembrane #status predicted <TM3>  
F:349-786/Domain: intracellular #status predicted <TM4>  
F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:787-810/Domain: transmembrane #status predicted <TM5>  
F:849-874/Domain: transmembrane #status predicted <TM6>  
F:875-952/Domain: intracellular #status predicted <INT4>  
F:953-978/Domain: transmembrane #status predicted <TM7>  
F:979-1023/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:508/Binding site: ATP (Lys) #status predicted  
F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 37.7%; Score 43; DB 1; Length 1023;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 84 PTPREWVFC 93

RESULT 42  
A60444  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain precursor - African clawed frog  
N:Alternate names: sodium pump alpha chain  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 09-Jul-2004  
C:Accession: A60444  
F:Verrey, F.; Kallouz, P.; Schaefer, E.; Fuentes, P.; Geering, K.; Rossier, B.C.; Kraet  
Am. J. Physiol. 256, F1034-F1043, 1989  
A:Title: Primary sequence of Xenopus laevis Na(+)-K(+)-ATPase and its localization in  
A:Reference number: A60444; MUID:89285429; PMID:2544104  
A:Accession: A60444  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1025 <VER>  
A:Cross-references: UNIPROT:Q92123; GB:U10108; NID:g499225; PIDN:AAA19022.1; PID:g49922  
C:Comment: The alpha chain is the catalytic chain.  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP, glycoprotein, hydrolase, phosphoprotein, potassium transport, sodium t  
F:589-785/Domain: intracellular #status predicted <INT3>  
F:217,485/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:510/Binding site: ATP (Lys) #status predicted

Query Match 37.7%; Score 43; DB 2; Length 1025;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 PTLREWISFC 16  
Db 86 PTPREWVFC 95

RESULT 43  
PWCMM  
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - white sucker  
C:Species: Catostomus commersoni (white sucker)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S14740  
R:Schoenrock, C.; Morley, S.D.; Okawara, Y.; Lederis, K.; Richter, D.  
Biol. Chem. Hoppe-Seyler 372, 279-286, 1991  
A:Title: Sodium and potassium ATPase of the teleost fish Catostomus commersoni. Sequen  
A:Reference number: S14740; MUID:91282983; PMID:1711856  
A:Accession: S14740  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1027 <SCH>  
A:Cross-references: UNIPROT:P25489; EMBL:X58629; NID:g62641; PIDN:CAA41483.1; PID:g6264  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP, hydrolase, ion transport, phosphoprotein, potassium transport, sodium  
F:99-124/Domain: transmembrane #status predicted <TM1>  
F:133-152/Domain: transmembrane #status predicted <TM2>  
F:153-293/Domain: intracellular #status predicted <INT2>  
F:294-316/Domain: transmembrane #status predicted <TM3>  
F:323-351/Domain: transmembrane #status predicted <TM4>  
F:352-790/Domain: intracellular #status predicted <INT3>  
F:591-787/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:791-814/Domain: transmembrane #status predicted <TM5>  
F:853-878/Domain: transmembrane #status predicted <TM6>  
F:879-956/Domain: intracellular #status predicted <INT4>  
F:957-982/Domain: transmembrane #status predicted <TM7>

F\_983-1027/Domain: extracellular #status predicted <EXT>  
F\_379/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F\_512/Binding site: ATP (lys) #status predicted  
F\_721,725,730/Active site: Asp, Asp, Lys #status predicted

Query Match	37.7%	Score 43;	DB 1;	Length 1027;
Best Local Similarity	60.0%	Pred. No. 2.6e+02;		
Matches	6;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0;

```
Qy      7 PTLREWISFC 16
          |||: ||
Db      87 PTPPEWKFC 96
```

**RESULT 44**

Na<sup>+</sup>/K<sup>+</sup>-exchanging ATPase (EC 3.6.3.9) alpha chain - fruit fly (*Drosophila melanogaster*)

N/Alternate names: sodium pump alpha chain  
C/Species: *Drosophila melanogaster*  
C/Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
C/Accession: S03632, S07049  
E/Leibovitz, R.M.; Takeyasu, K.; Fambrough, D.M.

EMBO J. 8, 193-202, 1989  
A:Title: Molecular characterization and expression of the (Na+K)-ATPase alpha-subunit in A:Reference number: S03632; MUID:89231618, PMID:2540956  
A:Accession: S03632  
A:Molecule type: mRNA

A:Residues: 1-1038 <LEB>  
A:Cross-references: UNIPROT:P13607; EMBL:X14476  
A:Note: the sequence from Fig. 9 is inconsistent with that from Fig. 8 in having 89-Asp,  
R,Valadi, A.; Gilmore-Heber, M.; Benz Jr., E.J.  
FEBS Lett. 258, 203-207, 1989  
A:Title: Amplification of the phosphorylation site - ATP-binding site cDNA fragment of t  
A:Reference number: S07049; MUID:90092469; PMID:2557235  
A:Accession: S07049  
A:Molecule type: mRNA  
A:Residues: 397-521 <VAR>  
A:Cross-references: EMBL:X17471  
A:Note: the authors translated the codon ACC for residue 3 as Asn and AAT for residue 89

A:Gene: FlyBase:Atcp-alpha  
A:Cross-references: FlyBase:FBgm0002921  
A:Map position: 3R 93B  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
K:Keywords: ATP, heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

Query Match	37.7%;	Score 43;	DB 1;	length 1038;
Best Local Similarity	44.4%;	Pred. NO. 2.6e+02;		
Matches	8;	Conservative	1;	Mismatches 3;
				Indels 6;
				Gaps 1;

```
QY      5 DGPTLR-----EWISFC 16
          ||| |
          ||: ||
Db      93 DGNLTPPKQTPREWVKFC 110
```

RESULT 45  
A42687

neurotrophin-4 precursor - human  
N:Alternate names: neurotrophin-5  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: AA2687; J05050

Proc. Natl. Acad. Sci. U.S.A. 89, 3060-3064, 1992  
 A1Title: Mammalian neurotrophin-4: structure, chromosomal localization, tissue distribution  
 A1Reference number: A42687, MUID:92212967, PMID:1313578

A:Accession: A4768/  
A:Molecule type: DNA  
A:Residues: 1-210 <1P>  
A:Cross-references: UNIPROT:P34130, GB:M66528, NID:G190264, PIDN:AAA60154.1, PID:G190264  
A:Note: Sequence extracted from NCBI backbone (NCBI:G93810, NCBI:P.93811)  
R:Berkemeier, L.R.; Winslow, J.W.; Kaplan, D.R.; Nikolic, K.; Goeddel, D.V.; Rosenthal  
Neuron 7, 857-866, 1991

A:Title: Neurotrophin-5: a novel neurotrophic factor that activates Trk and TrkB.  
A:Reference number: JH0503, MUID:92075279; PMID:11742028  
A:Accession: JH0503  
A:Status: nucleic acid sequence not shown  
A:Stature: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <BER>  
C:Comment: The neurotrophins stimulate autophosphorylation and transduce signals through  
C:GeneticDB: NTF5  
C:GeneticDB: NTF5

A:Gene: GDB:134723; OMIM:162662	
A:Cross-References: 19pter-19pter	
A:Map position: 19pter-19pter	
C:Superfamily: nerve growth factor beta chain	
C:Keywords: glycoprotein	
F:1-24/Domain: signal sequence #status predicted <SIG>	
F:25-80/Domain: propeptide #status predicted <PRO>	
F:81-210/Product: neurotrophin-4 #status predicted <NEU>	
F:76/Binding site: carbohydrate (Asn) (covalent) #status predicted	
Query Match Similarity 37.3%	Score 42.5; DB 2; Length 210;
Best local similarity 47.4%	Pred. No. 71;
Matches 9; Conservative 1; Mismatches 8; Indels 1; Gaps 1	

QY 1 GGCADGPTLREWISFCGK 19  
||| | | | |  
: | |  
Db 156 GGCR-GVDRRHWSCKAK 173

Search completed: September 1, 2005, 16:22:51  
Job time : 15.4892 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 70.6691 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-8

Perfect score: 114  
Sequence: 1 GGCGADGPTLRWISFCGSK 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprotc:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	56	49.1	297	2 Q7UQB4	Q7UQB4 rhodopsin
2	55.5	48.7	934	2 Q9NEX6	Q9NEX6 caenothabdi
3	51	44.7	386	1 ETR1_CANTR	Q8wzm3 candida tro
4	51	44.7	386	1 ETR2_CANTR	Q8wzm4 candida tro
5	50.5	44.3	387	2 Q98A97	Q98A97 rhizobium 1
6	50.5	44.3	389	2 Q8KJF9	Q8KJF9 rhizobium 1
7	49.5	43.4	405	2 Q9KIE9	Q9KIE9 streptomyc
8	49	43.0	245	2 Q9M060	Q9M060 arabidopsis
9	49	43.0	349	2 Q7V2B2	Q7V2B2 prochloroc
10	48	42.1	319	2 Q9RKM5	Q9RKM5 streptomyc
11	48	42.1	342	2 Q6VNM4	Q6VNM4 streptomyc
12	48	42.1	361	2 Q7J2M7	Q7J2M7 mycobacteri
13	48	42.1	1123	2 Q7QC63	Q7QC63 anophelies g
14	47.5	41.7	238	2 Q7ULR5	Q7ULR5 rhodopsin
15	47.5	41.7	283	2 Q82CW2	Q82CW2 streptomyc
16	47	41.2	94	2 Q6MX73	Q6MX73 azarococcus sp
17	47	41.2	129	2 Q8DHX7	Q8DHX7 synecococc
18	47	41.2	271	2 Q8SPB8	Q8SPB8 bradyrhizob
19	47	41.2	475	2 Q9UAT5	Q9UAT5 caenothabdi
20	47	41.2	821	2 Q966D4	Q966D4 caenothabdi
21	47	41.2	956	2 Q6CLJ9	Q6CLJ9 kluveromyc
22	47	41.2	1926	2 Q9Y8B3	Q9Y8B3 paracoccidi
23	46.5	40.8	166	2 Q6KGG9	Q6KGG9 bacterioph
24	46.5	40.8	426	2 Q89HDB	Q89HDB bradyrhizob
25	46	40.4	97	2 Q8FPCA	Q8FPCA corynebacte
26	46	40.4	117	2 Q7MV49	Q7MV49 porphyromon
27	46	40.4	159	2 Q8N852	Q8N852 homo sapien
28	46	40.4	162	2 Q63KH8	Q63KH8 burkholderi
29	46	40.4	196	2 Q7VWMS	Q7VWMS burkholderi
30	46	40.4	196	2 Q7W9K1	Q7W9K1 burkholderi
31	46	40.4	245	2 Q8GVPS	Q8GVPS oryza sativ

32	46	40.4	275	2 O13090	O13090 pleurodeles
33	46	40.4	277	1 IMPI_MOUSE	P70224 mus musculu
34	46	40.4	312	2 Q9ND00	Q9ND00 trypanosoma
35	46	40.4	347	2 Q7PBP6	Q7PBP6 anophelies g
36	46	40.4	403	2 Q88N02	Q88N02 pseudomonas
37	46	40.4	443	2 Q9P858	Q9P858 phaeosphaer
38	46	40.4	482	2 Q6A1R0	Q6A1R0 desulfotale
39	46	40.4	540	2 Q82L10	Q82L10 streptomyc
40	46	40.4	926	1 AASS_HUMAN	Q94975 homo sapien
41	46	40.4	1902	2 Q9Y878	Q9Y878 coccidioid
42	45.5	39.9	309	2 Q8XZNS	Q8XZNS ralsionia s
43	45.5	39.9	485	2 Q8SC10	Q8SC10 propionibac
44	45	39.5	108	2 Q7RUA5	Q7RUA5 neurospora
45	45	39.5	146	2 Q6ZTT4	Q6ZTT4 homo sapien
46	45	39.5	173	2 Q8C4M6	Q8C4M6 mus musculu
47	45	39.5	180	2 Q07291	Q07291 natronomona
48	45	39.5	201	2 Q75LB6	Q75LB6 oryza sativ
49	45	39.5	209	2 Q6N1X5	Q6N1X5 rhodopsin
50	45	39.5	290	2 Q88HF5	Q88HF5 pseudomonas
51	45	39.5	290	2 Q88JF5	Q88JF5 bradyrhizob
52	45	39.5	338	2 Q82CX1	Q82CX1 streptomyc
53	45	39.5	367	2 Q64BD6	Q64BD6 uncultured
54	45	39.5	379	2 Q7SXV0	Q7SXV0 brachydanio
55	45	39.5	385	2 Q7XMK0	Q7XMK0 oryza sativ
56	45	39.5	410	2 Q629V1	Q629V1 burkholderi
57	45	39.5	410	2 Q63HZ6	Q63HZ6 burkholderi
58	45	39.5	421	2 Q9XUV7	Q9XUV7 caenothabdi
59	45	39.5	499	1 MEP2_YEAST	Q9XUV7 saccharomyc
60	45	39.5	594	2 Q7SHC4	Q7SHC4 neurospora
61	45	39.5	721	2 Q6K4D2	Q6K4D2 oryza sativ
62	45	39.5	769	2 Q70804	Q70804 tt virus. 1
63	45	39.5	894	2 Q63UA1	Q63UA1 burkholderi
64	45	39.5	1134	2 Q8P378	Q8P378 xanthomonas
65	45	39.5	1335	1 RFOR_HUMAN	Q8A122 homo sapien
66	45	39.5	175	2 Q7XQ02	Q7XQ02 oryza sativ
67	44.5	39.0	248	2 Q7PXF4	Q7PXF4 anophelies g
68	44.5	39.0	271	2 Q72J02	Q72J02 thermus the
69	44.5	39.0	282	2 Q7QCK2	Q7QCK2 anophelies g
70	44.5	39.0	282	2 Q8AVB0	Q8AVB0 brachydanio
71	44.5	39.0	429	2 Q7QIK7	Q7QIK7 anophelies g
72	44.5	39.0	497	2 Q7QIB6	Q7QIB6 brachydanio
73	44.5	39.0	818	2 Q6PBA6	Q6PBA6 brachydanio
74	44.5	39.0	1067	2 Q6M209	Q6M209 aspergillus
75	44.5	39.0	1142	2 Q8IT12	Q8IT12 trypanosoma
76	44.5	39.0	1183	2 Q7NLS9	Q7NLS9 gloeobacter
77	44	38.6	173	2 Q6ZAD7	Q6ZAD7 oryza sativ
78	44	38.6	178	2 Q6QHD2	Q6QHD2 gallid herp
79	44	38.6	197	2 Q6RL14	Q6RL14 gallid herp
80	44	38.6	209	2 Q6R8A0	Q6R8A0 sodalis glo
81	44	38.6	209	2 Q7L059	Q7L059 streptomyc
82	44	38.6	210	2 Q69PA9	Q69PA9 oryza sativ
83	44	38.6	238	2 Q7QKA0	Q7QKA0 anophelies g
84	44	38.6	292	2 Q67642	Q67642 gallid herp
85	44	38.6	298	2 Q86653	Q86653 gallid herp
86	44	38.6	304	2 Q82RG2	Q82RG2 streptomyc
87	44	38.6	310	2 Q678H2	Q678H2 lymphocyeti
88	44	38.6	346	2 Q62030	Q62030 caenothabdi
89	44	38.6	371	2 Q9RRJ3	Q9RRJ3 deinococcus
90	44	38.6	404	2 Q7QF40	Q7QF40 anophelies g
91	44	38.6	425	2 Q8PD03	Q8PD03 xanthomonas
92	44	38.6	450	2 Q75211	Q75211 ashbya goss
93	44	38.6	490	2 Q04270	Q04270 chlamydomon
94	44	38.6	519	2 Q7Y1N9	Q7Y1N9 oryza sativ
95	44	38.6	524	2 Q66GJ0	Q66GJ0 arabidopsis
96	44	38.6	524	2 Q84W33	Q84W33 arabidopsis
97	44	38.6	526	2 Q9C868	Q9C868 arabidopsis
98	44	38.6	537	2 Q63M37	Q63M37 burkholderi
99	44	38.6	613	2 Q7Y7Z5	Q7Y7Z5 burkholderi
100	44	38.6	613	2 Q9VGR8	Q9VGR8 dirosophila

## ALIGNMENTS

```

RESULT 1
Q7UOE4 PRELIMINARY; PRT; 297 AA.
ID Q7UOE4;
AC Q7UOE4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Hypothetical protein.
OS OrderedLocustNames=RB6375;
Rhodopirellula ballica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heltmann K., Rabus R.,
RA Schleener H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
errata 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL, BX294144; CAD74759.1; -.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR PROSITE; PS00152; ATPase_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50829; GVF; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475f670f02c7859b CRC64;

Query Match 49.1%; Score 56; DB 2; Length 297;
Best Local Similarity 69.2%; Pred. No. 2.5;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCADGPTLRWIS 14
| | | | | | | | | |
Db 173 GPADGPTMKOMIS 185

RESULT 2
Q9NEX6 PRELIMINARY; PRT; 934 AA.
ID Q9NEX6;
AC Q9NEX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Y10588A.21.
GN ORFNames=Y10588A.21;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018 (1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL332876; CAC48140.1; -.
DR WormBase; WBGen00013679; Y10588A.21.
DR WormPep; Y10588A.21; CE25162.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.

```

```

KM Hypothetical protein.
SQ SEQUENCE 934 AA; 10485 MW; 5ED4E1D03DB06F24 CRC64;

Query Match 48.7%; Score 55.5; DB 2; Length 934;
Best Local Similarity 55.6%; Pred. No. 9.2;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 3 CADGPTLRW-ISFCGK 19
| | | | | | | | | |
Db 899 CVDGTRDWPVSFTGSE 916

RESULT 3
ETRL_CANTR STANDARD; PRT; 386 AA.
ID ETRL_CANTR;
AC O8WMZ3;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Enoyl-[acyl-carrier protein] reductase [NADPH, B-specific] 1,
DE mitochondrial precursor (EC 1.3.1.10).
GN Name=ETRL;
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
[1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 23-29, FUNCTION, SUBUNIT, AND
SUBCELLULAR LOCATION.
RC STRAIN=ATCC 20336;
RX MEDLINE=21400968; PubMed=11509667;
RX DOI=10.1128/MCB.21.18.6243-6253.2001;
RA Torcko J.M., Kolivranca K.T., Minalainen I.J., Yagi A.I., Schmitz W.,
RA Kastaniotis A.J., Altman T.T., Gurvitz A., Hiltunen J.K.;
RT "Candida tropicalis Erip and Saccharomyces cerevisiae Ybr026p
(Mrf1p), 2-enoyl thioester reductases essential for mitochondrial
respiratory competence."
RL Mol. Cell. Biol. 21:6243-6253 (2001).
[2]
RP SUBUNIT.
RC STRAIN=ATCC 20336;
RX PubMed=12890667; DOI=10.1074/jbc.M307664200;
RA Torcko J.M., Kolivranca K.T., Kastaniotis A.J., Altman T.T.,
RA Glunoff T., Ilyes M., Hartig A., Gurvitz A., Hiltunen J.K.;
RT "Candida tropicalis expresses two mitochondrial 2-enoyl thioester
reductases that are able to form both homodimers and heterodimers."
RL J. Biol. Chem. 278:41213-41220 (2003).
[3]
RX X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS), AND MUTAGENESIS OF TYR-79.
RX PubMed=12614607; DOI=10.1016/S0022-2836(03)00038-X;
RA Altman T.T., Torcko J.M., Van den Plas S., Sormunen R.T.,
RA Kastaniotis A.J., Wierenga R.K., Hiltunen J.K.;
RT "Structure-function analysis of enoyl thioester reductase involved in
mitochondrial maintenance."
RL J. Mol. Biol. 327:47-59 (2003).
CC -|- FUNCTION: Required for respiration and the maintenance of the
mitochondrial compartment. May have a role in the mitochondrial
synthesis of fatty acids.
CC -|- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
CC -|- SUBUNIT: Homodimer and heterodimer with etrl2.
CC -|- SUBCELLULAR LOCATION: Mitochondrion.
CC -|- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
family. Quinone oxidoreductase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: U94997; AAL55472.1; -.
DR PDB: 1GU7; X-ray; A/B=23-386.
DR PDB: 1GU7; X-ray; A/B=23-386.
DR PDB: 1GYR; X-ray; A/B/C=23-386.
DR InterPro: IPR002085; Adh zn family.
DR InterPro: IPR011032; GroES like.
DR Pfam: PF00107; ADH_zinc_N_1.
DR 3D-structure; Direct protein sequencing; Fatty acid biosynthesis;
KM Mitochondrion; NADP; Oxidoreductase; Transit peptide.
FT TRANSIT 1 22 Mitochondrion.
FT CHAIN 23 386 Enoyl-[acyl-carrier protein] reductase
FT NTPAGEN 79 79 [NADPH, B-specific] 1.
SQ SEQUENCE 386 AA; 42160 MW; FCBC174A240742D8 CRC64;
Y->N: 0.1% of catalytic activity.

Query Match 44.7%; Score 51; DB 1; Length 386;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 GPTIREMISFCGK 19
DB 254 GPTIKWIKOSGGE 267

RESULT 4
ID ETR2_CANTR STANDARD; PRT; 386 AA.
AC 08WZM4;
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Enoyl-[acyl-carrier protein] reductase [NADPH, B-specific] 2,
DE Mitochondrial precursor (EC 1.3.1.10).
GN Name=ETR2;
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitospotic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN RP SEQUENCE FROM N.A., FUNCTION, AND SUBUNIT.
RC STRAIN=ATCC 20336;
RA PubMed=12890667; DOI=10.1074/jbc.M307664200;
RA Toroko J.M., Kojivuranca K.T., Kascariotis A.J., Airenne T.T.,
RA Glumoff T., Iives M., Hartig A., Guryitz A., Hiltunen J.K.;
RT "Candida tropicalis expresses two mitochondrial 2-enoil thioester
RT reductases that are able to form both homodimers and heterodimers.";
RT J. Biol. Chem. 278:41213-41220(2003).
RN RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.11 ANGSTROMS).
RA Airenne T.T., Toroko J.M., Hiltunen J.K.;
RT "Crystal structure of enoyl thioester reductase 2.";
RT Submitted (JUN-2002) to the PDB data bank.
CC -1- FUNCTION: Required for respiration and the maintenance of the
CC mitochondrial compartment. May have a role in the mitochondrial
CC synthesis of fatty acids.
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADPH(+) = trans-
CC -2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
CC -1- SUBUNIT: Homodimer and heterodimer with ETR1.
CC -1- SUBCELLULAR LOCATION: Mitochondrion (By similarity).
CC -1- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family. Quinone oxidoreductase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94996; AAL55471.1; -.
DR PDB: 1HOK; X-ray; A/B=23-386.
DR InterPro: IPR002085; Adh_zn_family.

```

```

DR InterPro: IPR011032; GroES like.
DR Pfam: PF00107; ADH_zinc_N_1.
DR 3D-structure; Fatty acid biosynthesis; Mitochondrion; NADP;
KM Oxidoreductase; Transit peptide.
FT TRANSIT 1 22 Mitochondrion (Potential).
FT CHAIN 23 386 Enoyl-[acyl-carrier protein] reductase
FT SEQUENCE 386 AA; 42116 MW; 91ABE00831F0C2E8 CRC64;
[ NADPH, B-specific] 2.

Query Match 44.7%; Score 51; DB 1; Length 386;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 GPTIREMISFCGK 19
DB 254 GPTIKWIKOSGGE 267

RESULT 5
ID 098A97 PRELIMINARY; PRT; 387 AA.
AC 098A97;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE M16096 protein.
GN OrderedLocustNames=m16096;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAEF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RT DNA Res. 7:331-338(2000).
DR EMBL: AP003008; BAB52440.1; -.
DR HSSP; P77407; 1PQY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro: IPR003673; CAIB BAIF.
DR Pfam: PF02515; COA_transf_3; 1.
KM Complete proteome.
SQ SEQUENCE 387 AA; 42226 MW; 64643BEC8F25518 CRC64;

Query Match 44.3%; Score 50.5; DB 2; Length 387;
Best Local Similarity 42.9%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

QY 3 CADGPTL-----REWISFC 16
DB 237 CADGEVIFSVQNDREWVNC 257

RESULT 6
ID 08KJF9 PRELIMINARY; PRT; 389 AA.
AC 08KJF9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PUTATIVE RACEMASE/DEHYDRATASE PROTEIN.
GN Name=m1181;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RTA;
RX MEDLINE=21999272; PubMed=12003951;
RX DOI=10.1128/JB.184.11.3086-3095.2002;
RA Sullivan J.T., Trzciatkowski J.R., Crickbank R.W., Gouzy J.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rosbach U.,
RA Stuart G.S., Weaver J.E., Webby R.U., de Bruijn F.J., Ronson C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A.";
RT J. Bacteriol. 184:3086-3095(2002).
DR EMBL: AF672113; CAD31586.1; -.
DR HSSP: P77407; 1FOY.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR003673; CAIB_BAIF.
DR Pfam: PF02515; COA_transf_3; 1.
SQ SEQUENCE 389 AA; 42703 MW; 667BD2C96A7E5204 CRC64;

Query Match 44.3%; Score 50.5; DB 2; Length 389;
Best Local Similarity 42.9%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Qy 3 CADGPTL-----REWISFC 16
Db 243 CADGKEVTFVSQNDREWMVFC 263

RESULT 7
O9KIE9 PRELIMINARY; PRT; 405 AA.
AC O9KIE9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FkBE.
GN Name=FkBE;
OS Streptomyces hygroscopicus subsp. acromyceticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OC NCBI_TaxID=132248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20323220; PubMed=10863099; DOI=10.1016/S0378-1119(00)00171-2;
RA Wu K., Chung L., Revill W.P., Katz L., Reeves C.D.;
RT "The FK520 gene cluster of Streptomyces hygroscopicus var.
RT acromyceticus (ATCC 14891) contains genes for biosynthesis of unusual
RT polypeptide extender units.";
RT Gene 251:81-90(2000)
RL EMBL: AF235504; AAF6384.1; -.
DR HSSP: P77407; 1FOY.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR003673; CAIB_BAIF.
DR Pfam: PF02515; COA_transf_3; 1.
SQ SEQUENCE 405 AA; 43696 MW; DC2569DFC914AD6F CRC64;

Query Match 43.4%; Score 49.5; DB 2; Length 405;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 10; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

Qy 5 DGPTL-----REWISFC 17
Db 252 DGOTINIGLONERWASFCG 271

RESULT 8
O9M060 PRELIMINARY; PRT; 245 AA.
AC O9M060;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 6 (EIF-6)-like protein
DE (Ac3955620).

```

```

GN Name=Flit6_30; Synonyms=At3g55620;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmach E., Drzonek H., Ansoze W., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinzaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161667; CAB81587.1; -.
DR EMBL: BT006566; AAF75806.1; -.
DR EMBL: AY128351; AAM91554.1; -.
DR PIR: T47701; T47701.
DR HSSP: Q12522; 1G62.
DR GO: GO:0003743; P:translation initiation factor activity; IEA.
DR GO: GO:0006413; P:translational initiation; IEA.
DR InterPro: IPR002769; eIF6.
DR Pfam: PF01912; eIF-6; 1.
DR ProDom: PD006880; eIF6; 1.
DR SMART: SM00654; eIF6; 1.
DR TIGRFAMs: TIGR00323; eIF-6; 1.
DR KX
SQ SEQUENCE 245 AA; 26482 MW; 73369A2A657F390D CRC64;

Query Match 43.0%; Score 49; DB 2; Length 245;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ADGPTLREWISFCG 17
Db 194 AAGMTVNDWTSFCG 207

RESULT 9
O7V2B2 PRELIMINARY; PRT; 349 AA.
ID O7V2B2;
AC O7V2B2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dihydroorotase (EC 3.5.2.3).
GN Name=pyrC; OrderedLocNames=PMM0569;
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcales;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;

```

RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,  
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hens W.R.,  
 RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,  
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
 RA Webb E.A., Zinser E.R., Chisholm S.W.;  
 RT "Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic  
 niche differentiation";  
 RL Nature 424:1042-1047(2003).  
 DR EMBL; BX572091; CAE19028.1; -.  
 DR HSSP; P05020; 1079.  
 DR GO; GO:0004151; F:dihydroorotase activity; IEA.  
 DR GO; GO:0016787; P:hydrolase activity; IEA.  
 DR GO; GO:0019856; P:pyrimidine base biosynthesis; IEA.  
 DR InterPro; IPR006680; Amidohydro\_1.  
 DR InterPro; IPR004721; Dihodimr..  
 DR InterPro; IPR002195; Dihydroorotase.  
 DR Pfam; PF01979; Amidohydro\_1; 1.  
 DR TIGRFAms; TIGR00856; pyrC dimer; 1.  
 DR PROSITE; PS00482; DIHYDROOROTASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00483; DIHYDROOROTASE\_2; 1.  
 KM Complete proteome; Hydrolase.  
 SQ SEQUENCE 349 AA; 39958 MW; CC02F5AE02EC927 CRC64;  
 Query Match 43.0%; Score 49; DB 2; Length 349;  
 Best Local Similarity 50.0%; Pred. No. 35;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GCADGPTLRWISFCG 17  
 Db 243 GTDSAPHLRQWKAFCG 258

RESULT 10

Q9RKM5 PRELIMINARY; PRT; 319 AA.  
 AC Q9RKM5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Putative Meir family transcriptional regulator.  
 GN ORFNames=SCD17.06C;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(12) / M145;  
 RX MEDLINE=2196410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wetzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete *Streptomyces*  
 RT *coelicolor* A3(2)";  
 RL Nature 417:141-147(2002).  
 CC -1 - SIMILARITY: Contains 1 HTH mer-r-type DNA-binding domain.  
 DR EMBL; AL939118; CAB56383.1; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000551; HTH\_MerR.  
 DR InterPro; IPR009061; Putativ\_DNA\_bind.  
 DR Pfam; PF00376; MerR; 1.  
 DR PRINTS; PR00040; HTHMER.  
 DR SMART; SM00422; HTH\_MER\_1.  
 DR PROSITE; PS0037; HTH\_MER\_2; 1.  
 KM Complete proteome; DNA-binding.

SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;  
 Query Match 42.1%; Score 48; DB 2; Length 319;  
 Best Local Similarity 61.5%; Pred. No. 46;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GCADGPTLRWIS 14  
 Db 255 GRPDGPTLRWLA 267

RESULT 11

Q6VMH4 PRELIMINARY; PRT; 342 AA.  
 AC Q6VMH4;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Putative SARF family pathway specific regulatory protein.  
 GN Name=slpU;  
 OS Streptomyces ambofaciens.  
 OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=1889;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 23877;  
 RX PubMed=14742212;  
 RA Pang X., Aigle B., Girardet J.M., Mangenot S., Pernodet J.L.,  
 RA Decaris B., Leblond P.;  
 RT "Functional repeats of the Streptomyces ambofaciens linear chromosome";  
 RL Antimicrob. Agents Chemother. 48:575-588(2004).  
 DR EMBL; AY338477; AAN30165.1; -.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000156; P:two-component response regulator activity; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
 DR InterPro; IPR009059; b1\_resp\_regltr\_C.  
 DR InterPro; IPR005158; BTAD.  
 DR InterPro; IPR001867; Trans\_reg\_C.  
 DR Pfam; PF03704; BTAD; 1.  
 DR Pfam; PF00486; Trans\_reg\_C; 1.  
 SQ SEQUENCE 342 AA; 35639 MW; 945BC929E5ACEE3D CRC64;  
 Query Match 42.1%; Score 48; DB 2; Length 342;  
 Best Local Similarity 46.7%; Pred. No. 50;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GCADGPTLRWISFC 16  
 Db 112 GCGGPFSSRPLWBS 126

RESULT 12

Q73ZW7 PRELIMINARY; PRT; 461 AA.  
 AC Q73ZW7;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocNames=MAP1484C;  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K10;  
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;  
 RL EMBL; AE017232; AAS03801.1; -.  
 DR GO; GO:0005506; P:iron ion binding; IEA.

```

DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006725; P:aromatic compound metabolism; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR005806; R:Rieske reg.
DR InterPro: IPR001663; R:ring_hydroxyl_A.
DR Pfam: PF00355; R:Rieske_1.
DR PRINTS: PR00901; R:KMGDIOLINASE.
DR Complete proteome.
SQ SEQUENCE 461 AA; 52010 MW; 208E39A89C121839 CRC64;

Query Match 42.1%; Score 48; DB 2; Length 461;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Qy 1 GGCAC-----DGPFLREMI 13
    |||||
    |||||
Db 156 GGCACWNLDDAPALRMM 174

RESULT 13
ID 070C63 PRELIMINARY; PRT; 1123 AA.
AC 070C63;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP1221.
GN Name=agCG53078; ORFNames=ENSNAGC00000018865;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008859; EAA08177.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008898; F:homocysteine S-methyltransferase activity; IEA.
DR GO: GO:0004672; F:protein kinase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; K:kinase_like.
DR InterPro: IPR000719; P:prot_kinase.
DR InterPro: IPR003726; S:methyl_trans.
DR Pfam: PF00069; P:kinase_1.
DR Pfam: PF02574; S:methyl_trans_1.
DR ProDom: PD000001; P:prot_kinase_1.
DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.
SQ SEQUENCE 1123 AA; 12006 MW; D3CC001D8D4882AF CRC64;

Query Match 42.1%; Score 48; DB 2; Length 1123;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 ADGPFLREWISF 15
    |||||
    |||||
Db 969 ADHPVRFWISF 980

RESULT 14
ID 070UR5 PRELIMINARY; PRT; 238 AA.
AC 070UR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to phycoerythrin alpha phycoerythrin lyase Cgce (EC 4.-.-.-
DE ).
GN Name=cgce; OrderedLocNames=RB9340;

```

```

OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Firellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Hellmann K., Rabus R.,
RA Schlesner H., Aumann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Firellula sp.
RT strain 1."
RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL: BX294149; CAD76204.1; -.
DR GO: GO:0016829; F:lyase activity; IEA.
DR InterPro: IPR008938; ARM.
DR InterPro: IPR004155; ARM_lyase_HEAT.
DR Pfam: PF03130; HEAT_PBS_1.
DR SMART: SM00567; EZ_HEAT; 3.
DR Complete proteome; Lyase.
SQ SEQUENCE 238 AA; 26142 MW; B7CA7284593B0C72 CRC64;

Query Match 41.7%; Score 47.5; DB 2; Length 238;
Best Local Similarity 39.1%; Pred. No. 41;
Matches 9; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

Qy 1 GGCADSP-----TLREMI 12
    |||||
    |||||
Db 29 GGCADSPYALKHNPYFTRGW 51

RESULT 15
ID 082CW2 PRELIMINARY; PRT; 283 AA.
AC 082CW2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative ICLR-family transcriptional regulator.
GN OrderedLocNames=SAV5226;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL: AP005042; BAC72938.1; -.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR005471; HTH_ICLR.
DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
DR Pfam: PF01614; ICLR; 1.
DR Complete proteome.
SQ SEQUENCE 283 AA; 30503 MW; F63B1705578EE67 CRC64;

```

Query Match 41.7%; Score 47.5; DB 2; Length 283;  
Best Local Similarity 50.0%; Pred. No. 49;  
Matches 8; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 3 CADGPT--LREWISF 15  
|||:||||:|  
Db 152 CAGSPTTPAVHEWDF 167

## RESULT 16

Q6MX73 PRELIMINARY; PRT; 94 AA.

AC Q6MX73; 05-JUN-2004 (TREMBlrel. 27, Created)  
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=c2B002;  
OS Azarcus sp. (strain EBN1).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;  
OC Rhodocyclaceae; Azarcus.  
OX NCBI\_TaxID=76114;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EBN1;  
RA Kude M., Heider J., Amann J., Hufnagel P., Kuehner S., Beck A.,  
RA Reinhardt R., Rabus R.;  
RL Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EBN1;  
RA PROSCIENCE;  
RL Submitted (NCV-2003) to the EMBL/Genbank/DBJ databases.  
DR EMBL; BX682953; CAF21985.1; -.  
KW Hypothetical protein.

Qy SEQUENCE 94 AA; 9829 MW; 91AC4ECGABER7PDE CRC64;

Query Match 41.2%; Score 47; DB 2; Length 94;  
Best Local Similarity 72.7%; Pred. No. 20;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGCADGPTLRE 11  
|||:||||:|  
Db 46 GGCHDGAITRE 56

## RESULT 17

Q8DHX7 PRELIMINARY; PRT; 129 AA.

AC Q8DHX7; 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DE T111816 protein.  
GN OrderedLocusNames=c111816;  
OS Synechococcus elongatus (Thermosynechococcus elongatus).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
OX NCBI\_TaxID=32046;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BP-1;  
RX MEDLINE=22225144; PubMed=12240834;  
RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,  
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,  
RA Kiyokawa C., Kohara M., Takeuchi M., Matsuno M., Nakazaki N.,  
RA Shimo S., Sugimoto M., Matsuno M., Yamada M., Tabata S.;  
RT "Complete genome structure of the thermophilic cyanobacterium  
Thermosynechococcus elongatus BP-1."  
RL DNA Res. 9:123-130(2002).  
DR EMBL; AP005375; BAC09368.1; -.  
KW Complete proteome.  
SQ SEQUENCE 129 AA; 14644 MW; EBB44691E7DD1E12 CRC64;

Query Match 41.2%; Score 47; DB 2; Length 129;  
Best Local Similarity 56.2%; Pred. No. 27;  
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GCADGPTLREWISFCG 17  
|||:||||:|  
Db 38 GRAGATLRQWLSTLG 53

## RESULT 18

O89PE8 PRELIMINARY; PRT; 271 AA.

AC O89PE8; 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DE CufM protein.  
GN Name=cufM; OrderedLocusNames=blx3534;  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchimi T.,  
RA Sasamoto S., Watanabe A., Ideawa K., Iriyuchi M., Kawashima K.,  
RA Kohara M., Matsuno M., Shimo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110."  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AP005948; BAC48799.1; -.  
DR HSSP; P19920; 1N62.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR005107; CO\_den flav C.  
DR InterPro; IPR002346; dehydrog\_molyb.  
DR Pfam; PF00941; PAD\_binding\_5; 1.  
KW Complete proteome.  
SQ SEQUENCE 271 AA; 29422 MW; 4995C9FA814FDC6 CRC64;

Query Match 41.2%; Score 47; DB 2; Length 271;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCADGPTLREW 12  
|||:||||:|  
Db 211 GGADVPVPTARDW 222

## RESULT 19

O9UATS PRELIMINARY; PRT; 475 AA.

AC O9UATS; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE Hypothetical protein C01B4.7.  
GN Name=C01B4.7; ORFNames=C01B4.7;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";



```

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Smith A., Mamsley P., Fronick W.;
RT "The sequence of C. elegans cosmid C01B4.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF125952; AAD14699.1; -.
DR PIR; T33943; T33943.
DR Wormbase; WBGene0015271; C01B4.7.
DR WormPep; C01B4.7; CE20476.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
DR Hypothetical protein.
SQ SEQUENCE 475 AA; 53094 MW; 79095D45572AF535 CRC64;

Query Match 41.2%; Score 47; DB 2; Length 475;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCGG 18
DB 268 CTDRCVLSAWVSFLGG 283

RESULT 20
O966D4 PRELIMINARY; PRT; 821 AA.
AC O966D4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y19D10A.4.
DE Name=Y19D10A.4; ORFNames=Y19D10A.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG Wormbase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Johnson D.;
RT "The sequence of C. elegans cosmid Y19D10A.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```

```

RA Waterston R.H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006722; AAK68417.1; -.
DR PDB; 1LUR; X-ray; A/B=483-821.
DR Wormbase; WBGene00021219; Y19D10A.4.
DR WormPep; Y19D10A.4; CE21450.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004034; F:aldose 1-epimerase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR008183; Ald1_epimerase.
DR InterPro; IPR011013; Gal_mut_like.
DR InterPro; IPR007114; MFS.
DR Pfam; PF01263; Aldose_epim; 1.
DR PROSITE; PS50850; MFS; 1.
DR Hypothetical protein.
SQ SEQUENCE 821 AA; 91593 MW; 923A788FC95D1A76 CRC64;

Query Match 41.2%; Score 47; DB 2; Length 821;
Best Local Similarity 50.0%; Pred. No. 17+02;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFCGG 18
DB 268 CTDRCVLSAWVSFLGG 283

RESULT 21
O6CLJ9 PRELIMINARY; PRT; 956 AA.
ID O6CLJ9;
AC O6CLJ9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P40825|Saccharomyces cerevisiae YOR333c ALA1 alanyl-tRNA
DE synthetase.
OS ORFNames=KLA0F024319;
GN Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durieux P., Casaregola S.,
Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,

```



RA Goffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Blyskasten C.,  
 RA Boistrume A., Boyer J., Cattelico L., Confaitolieri F., de Daruvar A.,  
 RA Deepons L., Fabre E., Fairhead C., Ferry-Dumas H., Groppi A.,  
 RA Hantreay F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerret A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Micaud J.M., Nikolaki M., Ozias S., Ozler-Kalogeropoulos O.,  
 RA Pelenz S., Potter S., Richard G.F., Straub M.L., Suleau A.,  
 RA Sennene D., Tekala F., Wesolowski-Louvel M., Weschof E., Wirth B.,  
 RA Zenilou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchler C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
 RA Mincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382126; CAG97897.1;..  
 DR GO; GO:0004813; F:alanine-tRNA ligase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0006419; P:alanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR003156; Pesterase DHHA1.  
 DR InterPro; IPR002318; tRNA-synt\_2c.  
 DR InterPro; IPR006193; tRNA-synt\_Ala.  
 DR Pfam; PF02272; DHHA1; 1.  
 DR Pfam; PF01411; tRNA-synt\_2c; 1.  
 DR PRINTS; PRO0980; TRNASYNTHALA.  
 DR TIGRFAMs; TIGR00344; alas; 1.  
 DR PROSITE; PS50860; AA\_TRNA\_LIGASE\_I1\_ALA; 1.  
 DR Aminoacyl-tRNA synthetase.  
 KW SEQUENCE 956 AA; 107100 MW; 4F5CE6855880A3C CRC64;  
 SQ  
 Query Match 41.2%; Score 47; DB 2; Length 956;  
 Best Local Similarity 50.0%; Pred. No. 2e+02;  
 Matches 9; Conservative 1; Mismatches 4; Indels 4; Gaps 1;  
 Qy 5 DGPTLRW---ISFCGG 18  
 Db 704 ENPTSEWQKXISFCGG 721  
 RESULT 22  
 Q9Y8B3 PRELIMINARY; PRT; 1926 AA.  
 AC Q9Y8B3;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Glucan synthase.  
 GN Name=Fks;  
 OS Paracoccidioides brasiliensis.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Onygenales; mitosporic Onygenales; Paracoccidioides.  
 CX NCBI\_TaxID=121759;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PB01;  
 RX MEDLINE=20171859; PubMed=10705373;  
 RX DOI=10.1002/(SICI)1097-0061(20000330)16:5<451::AID-YEAS40>3.0.CO;2-O;  
 RA Pereira M., Felipe M.S.S., Brigido M.M., Soares C.M.A., Azevedo M.O.;  
 RT "Molecular cloning and characterization of a glucan synthase gene from  
 RT the human pathogenic fungus Paracoccidioides brasiliensis.";  
 RL Yeast 16:451-462(2000).  
 DR EMBL; AF148715; AAD37783.1;..  
 DR GO; GO:0000148; C:1,3-beta-glucan synthase complex; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0003843; F:1,3-beta-glucan synthase activity; IEA.  
 DR GO; GO:0006075; P:beta-1,3 glucan biosynthesis; IEA.  
 DR InterPro; IPR003440; Glyco\_trans\_48.  
 DR InterPro; IPR002114; HPT\_Serp\_S.

DR Pfam; PF02364; Glucan synthase; 1.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
 SQ SEQUENCE 1926 AA; 220574 MW; BB098950FP2253DS CRC64;  
 Query Match 41.2%; Score 47; DB 2; Length 1926;  
 Best Local Similarity 46.7%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 2 GCADPTLRWISFC 16  
 Db 1374 GCADPTLRWISFC 1388  
 RESULT 23  
 Q6KG99 PRELIMINARY; PRT; 166 AA.  
 AC Q6KG99;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Bacteriophage Felix 01.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.  
 CX NCBI\_TaxID=77775;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strangnathan N., Whitchard J.M., Pierson F.W., Kapur V., Weigt L.A.;  
 RT "Bacteriophage Felix 01: Genetic Characterization.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF320576; AAQ14824.1;..  
 KW Hypothetical protein.  
 KW SEQUENCE 166 AA; 19296 MW; 5AAB33E39DC3C989 CRC64;  
 SQ  
 Query Match 40.8%; Score 46.5; DB 2; Length 166;  
 Best Local Similarity 52.9%; Pred. No. 42;  
 Matches 9; Conservative 0; Mismatches 5; Indels 3; Gaps 1;  
 Qy 1 GGCADPTLRWISFC 17  
 Db 8 GSC---PTYGHWISLCG 21  
 RESULT 24  
 Q89HD8 PRELIMINARY; PRT; 426 AA.  
 AC Q89HD8;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE B1r603 protein.  
 GN OrderedNames=b1r603;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 CX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA110;  
 RX MEDLINE=2248498; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Idegawa K., Iriuchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Tsunooka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; AP005957; BAC1318.1;..  
 DR HSSP; P27017; 1Q00.  
 KW Complete proteome.  
 SQ SEQUENCE 426 AA; 47042 MW; AE20A1EC6CEB038 CRC64;  
 Query Match 40.8%; Score 46.5; DB 2; Length 426;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 GGCADGPTLRW 12  
 ||||: ||: |||  
 Db 416 GGCAG-PTFKW 426

## RESULT 25

Q8FPC4 PRELIMINARY; PRT; 97 AA.

AC Q8FPC4; 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=CE1858;  
 OS Corynebacterium efficiens;  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=152794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YC-314;  
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;  
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
 RA Gotohori T.;  
 RT "Comparative complete genome sequence analysis of the amino acid  
 RT replacements responsible for the thermostability of Corynebacterium  
 RT efficiens.";  
 RL Genome Res. 13:1572-1579(2003).  
 DR EMBL; AP005220; BAC18668.1; -  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 97 AA; 10632 MW; 6CPI1DA566B304C CRC64;

QY Query Match 40.4%; Score 46; DB 2; Length 97;  
 Best Local Similarity 58.3%; Pred. No. 29;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 77 GGCADGPTLRW 12  
 ||||: ||: |||

## RESULT 26

Q7MW49 PRELIMINARY; PRT; 117 AA.

AC Q7MW49; 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=P01251;  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Porphyromonadaceae; Porphyromonas.  
 OX NCBI\_TaxID=837;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W83;  
 RX MEDLINE=22829867; PubMed=12949112;  
 DOI=10.1128/JB.185.18.5591-5601.2003;  
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,  
 RA Eissen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,  
 RA Holt D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,  
 RA Dewhirst F.E., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,  
 RA Granger D., Tettelin H., Fraser C.M.;  
 RT "Complete genome sequence of the oral pathogenic bacterium  
 RT Porphyromonas gingivalis strain W83.";  
 RL J. Bacteriol. 185:5591-5601(2003).  
 DR EMBL; AB017176; AA066334.1; -  
 RN TIGR; PG1251; -  
 KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 117 AA; 12589 MW; B4421EB01D18689 CRC64;

QY Query Match 40.4%; Score 46; DB 2; Length 117;  
 Best Local Similarity 52.9%; Pred. No. 35;  
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Db 17 GGCYCVPTVAWIIIGAG 33  
 ||||: ||: |||

## RESULT 27

Q8N852 PRELIMINARY; PRT; 159 AA.

AC Q8N852; 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DE Hypothetical protein FLJ40008.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stomach;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
 RA Sudo H., Hoshiro T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Ninomiya K., Itohishi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tanai H., Kimura M., Watanabe M., Hirooka S., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T., Kusano J.,  
 RA Kanehori K., Takahashi-Fujii A., Hara R., Tanase T., Nomura Y.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshioka K., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Oono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohnori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 DR EMBL; AK097327; BAC04999.1; -  
 SQ SEQUENCE 159 AA; 17782 MW; DF63A4A6D73129A8 CRC64;

QY Query Match 40.4%; Score 46; DB 2; Length 159;  
 Best Local Similarity 52.6%; Pred. No. 48;  
 Matches 10; Conservative 0; Mismatches 5; Indels 4; Gaps 1;

Db 17 GGCGLIVKHGHTLRWNSF 35  
 ||||: ||: |||

## RESULT 28

Q63KH8 PRELIMINARY; PRT; 162 AA.

```

AC 063KB;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BPSS1183;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Tibball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bertley S.D., Sebailia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosser B., Davis P., Deshaizer D.,
RA Feltwell T., Frazer A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchaveest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; BX571966; CAH38855.1; -.
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 17186 MW; 27CDF4999112AB3 CRC64;

Query Match
Best Local Similarity 40.4%; Score 46; DB 2; Length 162;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GCADGPTLR 10
Db 93 GCADGPTLR 101

RESULT 29
Q7VWMS PRELIMINARY; PRT; 196 AA.
ID Q7VWMS;
AC Q7VWMS;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocustNames=BP2072;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leithner S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Umwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40 (2003).
DR EMBL; BX640428; CAH37057.1; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 196 AA; 21562 MW; D082FBA6A6C3A765 CRC64;

Query Match
Best Local Similarity 40.4%; Score 46; DB 2; Length 166;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GCADGPTLR 10
Db 18 GCASGPTLR 26

RESULT 31
Q8GVFS PRELIMINARY; PRT; 245 AA.
ID Q8GVFS;
AC Q8GVFS;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 26, Last annotation update)
DE Putative eukaryotic translation initiation factor 6.
GN Name=OU1340_C08.131;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone-OJ1340 C08."
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

```

```

Query Match
Best Local Similarity 40.4%; Score 46; DB 2; Length 196;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GCADGPTLR 10
Db 18 GCASGPTLR 26

RESULT 30
Q7W9KL PRELIMINARY; PRT; 196 AA.
ID Q7W9KL;
AC Q7W9KL;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocustNames=BP1756;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leithner S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Umwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40 (2003).
DR EMBL; BX640428; CAH37057.1; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 196 AA; 21562 MW; D082FBA6A6C3A765 CRC64;

Query Match
Best Local Similarity 40.4%; Score 46; DB 2; Length 196;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GCADGPTLR 10
Db 18 GCASGPTLR 26

RESULT 31
Q8GVFS PRELIMINARY; PRT; 245 AA.
ID Q8GVFS;
AC Q8GVFS;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 26, Last annotation update)
DE Putative eukaryotic translation initiation factor 6.
GN Name=OU1340_C08.131;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone-OJ1340 C08."
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

```



Db 144 CTDRALRDVACGCR 160

## RESULT 34

Q9NDD0 PRELIMINARY; PRT; 312 AA.

AC Q9NDD0; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)  
 DE Casein kinase 1 homolog 1 (Casein kinase 1.1).  
 GN Name=CKI.1;  
 OS Trypanosoma cruzi.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 NC NCB1\_TaxID=5693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Spadatora C., Repetto Y., Robello C., Morello A., Castany S.,  
 RA Gamarro F.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22276527; PubMed=12387847; DOI=10.1016/S0166-6851(02)00156-1;  
 RA Spadatora C., Repetto Y., Torres C., Pino L., Robello C., Morello A.,  
 RA Gamarro F., Castany S.;  
 RT "Two casein kinase 1 isoforms are differentially expressed in  
 Trypanosoma cruzi."  
 RL Mol. Biochem. Parasitol. 124:23-36 (2002).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AF164116; AAF80492.1; -.  
 DR EMBL; AF274060; AAK58697.1; -.  
 DR HSSP; 006486; 1CKI.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase\_1like.  
 DR InterPro; IPR000719; Prot Kinase.  
 DR InterPro; IPR008271; Ser\_Thr\_kin\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 312 AA; 35770 MW; 471E0BC2B0546321 CRC64;

Query Match 40.4%; Score 46; DB 2; Length 312;  
 Best Local Similarity 57.1%; Pred. No. 93;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

## RESULT 35

Q7PP6 PRELIMINARY; PRT; 347 AA.

AC Q7PP6; 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE ENSANGP00000020769 (Fragment).  
 GN Name=ENSANG00000018280;  
 OS Anopheles gambiae str. PEEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culioidae; Anopheles.  
 NC NCB1\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEEST;

RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.

DR EMBL; AAB01008944; EAA10075.2; -.  
 DR GO; GO:0008898; F:homocysteine S-methyltransferase activity; IEA.  
 DR InterPro; IPR003726; S\_methyl\_trans.  
 DR Pfam; PF02574; S-methyl\_trans; 1.  
 FT NON\_TER  
 SQ SEQUENCE 347 AA; 38585 MW; 66FF58A1000CDA4F CRC64;

Query Match 40.4%; Score 46; DB 2; Length 347;  
 Best Local Similarity 61.5%; Pred. No. 1e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CADGPTLRWISF 15  
 Db 201 CDEYPTVRFWISF 213

## RESULT 36

Q88NU2 PRELIMINARY; PRT; 403 AA.

AC Q88NU2; 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=PI1112;  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NC NCB1\_TaxID=160488;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22421060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.F.,  
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,  
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,  
 RA Moazzar A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,  
 RA Wedler H., Luder J., Scjepandic D., Hohelsel J., Straetz M., Heim S.,  
 RA Klewitz C., Bisen J.A., Tilmis K.N., Duesterhoeft A., Tuenmler B.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 metabolically versatile Pseudomonas putida KT2440."  
 RL Environ. Microbiol. 4:799-806 (2002).  
 DR EMBL; AE016778; AAN6737.1; -.  
 DR TIGR; PP112; -.  
 KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 403 AA; 42380 MW; 4D71AA1F370C58A7 CRC64;

Query Match 40.4%; Score 46; DB 2; Length 403;  
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GPTLRWISFCG 18  
 Db 149 GPTLRWLRDVCG 161

## RESULT 37

Q9P858 PRELIMINARY; PRT; 443 AA.

AC Q9P858; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Phaeosphaeria nodorum (Septoria nodorum).  
 OC Plasmid plasm1.

```

OC Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=13684;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BS444;
RA Rawson J.M.;
RT "Transposable elements in the phytopathogenic fungus Stagonospora
nodorum.";
RL Thesis (2000), PhD thesis, University of Birmingham, UK.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BS444;
RA Rawson J.M.; Cutler S.B., Caten C.E.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277966; CAB91876.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 443 AA; 4946 MW; 367E0762EB839568 CRC64;

Query Match 40.4%; Score 46; DB 2; Length 443;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CADGPTREWISPCGG 18
   ||| ||| |||
Db 170 CSENGTLEWITALQG 185

RESULT 38
ID O6A1T0 PRELIMINARY; PRT; 482 AA.
AC O6A1T0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=DP3021;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
  Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
  Tseling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
  Kleink H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
  from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:987-902(2004).
DR EMBL; CR522870; CAG3750.1; -.
DR InterPro; IPR003846; UPF0061.
DR Pfam; PF02696; UPF0061.1.
KW Complete proteome.
SQ SEQUENCE 482 AA; 54161 MW; 5F401BE29D89323D CRC64;

Query Match 40.4%; Score 46; DB 2; Length 482;
Best Local Similarity 69.2%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCADGPTLEWMI 13
   ||| ||| |||
Db 120 GRCAVGPALREFI 132

RESULT 39
ID O82L10 PRELIMINARY; PRT; 540 AA.
AC O82L10;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative long chain-fatty acid CoA ligase.
GN OrderedLocustNames=SAV2030;
OC Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
  Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
  Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
  avermitilis: deducing the ability of producing secondary
  metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
  Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
  microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
  family.
DR EMBL; AF005029; BAC69741.1; -.
DR HSSP; P08659; ILCI.
DR GO; GO:0016874; F.ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP BINDING; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 540 AA; 58879 MW; B3FBF500B20FC64 CRC64;

Query Match 40.4%; Score 46; DB 2; Length 540;
Best Local Similarity 56.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ADGPTREWISPCGK 19
   ||| ||| |||
Db 485 ADGPTLEWMAFCGQ 500

RESULT 40
ID AAS5_HUMAN STANDARD; PRT; 926 AA.
AC Q9UDR5; O95462;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-aminoacidic semialdehyde synthase, mitochondrial precursor
  (LKR/SDH) [includes: Lysine ketoglutarate reductase (EC 1.5.1.8) (LOR)
  (LKR); Saccharopine dehydrogenase (EC 1.5.1.9) (SDH)].
GN Name=AAS5;
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX PubMed=10775527;
RA Sacksteder K.A., Biery B.J., Morrell J.C., Goodman B.K.,
  Geisbrecht B.V., Cox R.P., Gould S.J., Geraghty M.T.;
RT "Identification of the alpha-aminoacidic semialdehyde synthase gene,
  which is defective in familial hyperlysinemia.";
RL Am. J. Hum. Genet. 66:1736-1743(2000).
RN [2]

```

RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Papes F., Kemper E.L., Cord-Neto G., Langone F., Arruda P.;  
 RT "Cloning and expression analysis of the LXR/SDH gene in human  
 tissues";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;  
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,  
 Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,  
 Ray J.L., Sekhon M., Becker M.C., Olafsholm M.D., Schaller M.E., Du H.,  
 Rens H., Edwards J., Bradshaw-Cordum H., All J., Andrews S., Isak A.,  
 Vabnick P., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,  
 Ozerly P., Bielicki L., Scott K., Holmes A., Harting R., Harris A.,  
 Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,  
 Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,  
 Tin-William A.-M., Abbott A., Mink P., Maupin R., Strommatt C.,  
 Latreille P., Miller N., Johnson D., Murray J., Moesner J.P.,  
 Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,  
 Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,  
 Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,  
 Clifton S.W., Chissole S.L., Marra M.A., Raymond C., Haugen E.,  
 Gillet W., Zhou Y., James R., Phelps K., Iadonoto S., Bubb K.,  
 Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Flurey T.S.,  
 Baerbach R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,  
 Bailey S.R., McPherson J.D., Olson M.V., Eichler E.B., Green E.D.,  
 Waterston R.H., Wilson R.K.;  
 RT "The DNA sequence of human chromosome 7";  
 RL Nature 424:157-164(2003).  
 CC -1- FUNCTION: A bifunctional enzyme that catalyzes the first two steps  
 in lysine degradation. The N-terminal and the C-terminal contain  
 lysine-ketoglutarate reductase and saccharopine dehydrogenase  
 activity, respectively.  
 CC -1- CATALYTIC ACTIVITY: N(6)-(L-1,3-dicarboxypropyl)-L-lysine +  
 NADP(+) + H(2)O = L-lysine + 2-oxoglutarate + NADPH.  
 CC -1- CATALYTIC ACTIVITY: N(6)-(L-1,3-dicarboxypropyl)-L-lysine + NMD(+) +  
 H(2)O = L-glutamate + 2-aminoadipate 6-semialdehyde + NADH.  
 CC -1- PATHWAY: lysine degradation; Saccharopine pathway; first-step.  
 CC -1- PATHWAY: lysine degradation; Saccharopine pathway; second step.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed in all 16 tissues examined with  
 highest expression in the liver.  
 CC -1- INDUCTION: Induced by starvation (By similarity).  
 CC -1- DISEASE: Defects in AAS are the cause of hyperlysinemia  
 (MIM:338700). Hyperlysinemia is an autosomal recessive condition  
 characterized by hyperlysinemia, lysinuria and variable  
 saccharopinuria.  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the ALADH/PNT  
 family.  
 CC -1- SIMILARITY: In the C-terminal section; belongs to the saccharopine  
 dehydrogenase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AF239180; AAF44328.1; -  
 DR EMBL; AJ007714; CAA07619.2; -  
 DR EMBL; AC006020; AAF03526.1; -  
 DR GeneW; HGNC:17366; AAS.  
 DR Reactome; Q9UDR5; -  
 DR MIM; 605113; -  
 DR MIM; 238700; -  
 DR InterPro; IPR007698; ALADH\_PNT\_C.  
 DR InterPro; IPR007886; ALADH\_PNT\_N.

DR InterPro; IPR005097; Saccharop-dh.  
 DR Pfam; PF01262; ALADH\_PNT\_C; 1.  
 DR Pfam; PF05222; ALADH\_PNT\_N; 1.  
 DR Pfam; PF03435; Saccharop-dh; 1.  
 KW Mitochondrion; Multifunctional enzyme; NAD; NADP; Oxidoreductase;  
 KW Transit Peptide.  
 FT TRANSIT 1 32 Mitochondrion (By similarity).  
 FT CHAIN 33 926 Alpha-aminoacidic semialdehyde synthase.  
 FT DOMAIN 33 455 Lysine-ketoglutarate reductase.  
 FT DOMAIN 477 926 Saccharopine dehydrogenase.  
 FT CONFLICT 589 589 S -> C (in Ref. 2).  
 SQ SEQUENCE 926 AA; 102131 MW; CB4194014351A18D CRC64;  
 QY Query Match 40.4%; Score 46; DB 1; Length 926;  
 DB Best Local Similarity 53.8%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 3; Mismatches 3; Gaps 0;  
 6 GPTLRWISFRCGG 18  
 623 GATIBSTIYICGG 635  
 RESULT 41  
 ID 09Y878 PRELIMINARY; PRT; 1902 AA.  
 AC 09Y878;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Glucan synthase.  
 GN Name=FKS1;  
 OS Coccidioides posadasii.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Oryziales; mitosporic Oryziales; Coccidioides.  
 OX NCBI\_TaxID=199306;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Silveira;  
 RA Siegel E.M., Orsborn K.I., Galgiani J.N.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF159533; AAD45326.2; -  
 DR GO; GO:000148; C:1,3-beta-glucan synthase complex; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0003843; F:1,3-beta-glucan synthase activity; IEA.  
 DR GO; GO:0006075; P:beta-1,3 glucan biosynthesis; IEA.  
 DR InterPro; IPR003440; Glyco-trans. 48.  
 DR Pfam; PF02364; Glucan synthase; 1.  
 SQ SEQUENCE 1902 AA; 217552 MW; 66FC3C60E725F2F CRC64;  
 QY Query Match 40.4%; Score 46; DB 2; Length 1902;  
 DB Best Local Similarity 46.7%; Pred. No. 5.5e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 2 GCADGPTLRWISFC 16  
 1381 GCADINPRDWMVQRC 1395  
 RESULT 42  
 ID 08XZNS PRELIMINARY; PRT; 309 AA.  
 AC 08XZNS;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE PROBABLE TRANSCRIPTION REGULATOR PROTEIN.  
 GN Name=RS04642; Order=deocucNames=RS03360;  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```

RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Margenot S.,
RA Arlat M., Billault A., Brociter P., Camus J.C., Catolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cumac S., Demange N.,
RA Gaspin C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646064; CADI5062.1; -.
DR HSSP; Q9KXC7; I1XC.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 33774 MW; 733551741CE83182 CRC64;

Query Match 39.9%; Score 45.5; DB 2; Length 309;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1 GG--CADGPTLR-----EW--ISFCGK 19
DB 216 GGTMECTDGAVALREW 230

RESULT 43
Q88CT0 PRELIMINARY; PRT; 485 AA.
AC Q88CT0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Orl6.
OS Propionibacterium phage phiB5.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=189836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2186396; PubMed=1189111;
RX DOI=10.1128/JB.184.7.2030-2033.2002;
RA Chopin M.C., Rouault A., Ehrlich S.D., Gautier M.;
RT "Phlebotomous phage active on the gram-positive bacterium
RT Propionibacterium freudenreichii.";
RL U. Bacteriol. 184:2030-2033(2002).
DR EMBL; AF428260; AAL91699.1; -.
SQ SEQUENCE 485 AA; 48825 MW; 0B4F44ABE3DE91A4 CRC64;

Query Match 39.9%; Score 45.5; DB 2; Length 485;
Best Local Similarity 39.3%; Pred. No. 1.7e+02;
Matches 11; Conservative 4; Mismatches 4; Indels 9; Gaps 3;

QY 1 GG--CADGPTLR-----EW--ISFCGK 19
DB 418 GGAECGGGPTINLPAGVSWRLSWCGE 445

RESULT 44
Q7RUA5 PRELIMINARY; PRT; 108 AA.
AC Q7RUA5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein B24B19.30.
GN Name=NCU03933.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehm A.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selltremlkoef C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kotle G.O., Ueda G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gierre S.,
RA Kamal M., Kamyssele M., Mauceli B., Bielke C., Rudd S., Fishman D.,
RA Kytrofova S., Raasmussen C., Metzenberg R.L., Perkins D.D., Kroen S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmati S.A.,
RA Desonza C.C., Glass L., Orbach M.U., Berlund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebdole D.J., Freltag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
RU Nature 0:0-0(2003).
CC -1- CAUTION: the sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000719; BAA28336.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11994 MW; 093DC0D9617A252E CRC64;

Query Match 39.5%; Score 45; DB 2; Length 108;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 CADGPTLRWISFC 16
DB 70 CQCQPTLRWLSWC 83

RESULT 45
Q6ZTT4 PRELIMINARY; PRT; 146 AA.
AC Q6ZTT4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ44235.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Houta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Iishi S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Nagatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Negai K., Isogai T.,
RA Submlrted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AKI26223; BAC86495.1; -.
SQ SEQUENCE 146 AA; 16475 MW; C0B7BBE49151B89B CRC64;

Query Match 39.5%; Score 45; DB 2; Length 146;
Best Local Similarity 69.2%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCADGPTLRWIS 14
DB 28 GCADGCVLRQYIS 40

Search completed: September 1, 2005, 16:21:02
Job time : 71.6691 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 82.7482 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-9  
Perfect score: 97  
Sequence: 1 TIKPTLRQWLKSRHTS 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	18	2	AaW09499 Thrombopo
2	97	100.0	18	2	AaW09459 Thrombopo
3	97	100.0	18	2	AaW36650 Thrombopo
4	97	100.0	18	2	AaW33026 Thrombopo
5	97	100.0	18	3	AaB17024 TPO-mimet
6	97	100.0	18	4	AaU25869 Human thr
7	97	100.0	18	4	AaU25823 Human thr
8	97	100.0	18	5	AaB72910 TPO mimet
9	97	100.0	18	7	AdJ73062 TPO mimet
10	97	100.0	18	8	AdJ52697 CHI delet
11	97	100.0	18	8	AdJ51658 CHI delet
12	97	100.0	18	8	AaW09491 Thrombopo
13	97	100.0	18	2	AaW09491 Thrombopo
14	97	100.0	18	2	AaW35418 Thrombopo
15	97	100.0	18	2	AaW36642 Thrombopo
16	97	100.0	18	2	AaU25861 Human thr
17	97	100.0	18	4	AaU25998 Human thr
18	97	100.0	18	5	Abp51693 TPO mimet
19	97	100.0	18	5	Abp51691 TPO mimet
20	97	100.0	18	8	AdQ16625 TPO mimet
21	97	100.0	18	8	AdQ16629 TPO mimet
22	97	100.0	18	2	AaW09493 Thrombopo
23	97	100.0	18	2	AaW36644 Thrombopo
24	97	100.0	18	2	AaU25863 Human thr
25	97	100.0	18	4	Abg71748 Anticbody
					AaU26006 Human thr

26	58	59.8	18	2	AaW09460
27	58	59.8	18	2	AaW09498
28	58	59.8	18	2	AaW36649
29	58	59.8	18	2	AaW33027
30	58	59.8	18	2	AaW36652
31	58	59.8	18	3	AaB17026
32	58	59.8	18	4	AaU25868
33	58	59.8	18	4	AaU25824
34	58	59.8	18	4	AaU25871
35	58	59.8	18	5	AbB72912
36	58	59.8	18	7	AdJ73064
37	58	59.8	18	8	AdJ52699
38	58	59.8	18	8	AdJ51660
39	58	59.8	128	8	AdQ16705
40	58	59.8	225	8	AdQ16704
41	57	58.8	13	2	AaW36779
42	57	58.8	14	2	AaW09463
43	57	58.8	14	2	AaW09468
44	57	58.8	14	2	AaW33030
45	57	58.8	14	2	AaW33034
46	57	58.8	14	2	AaW36774
47	57	58.8	14	2	AdJ24843
48	57	58.8	14	3	AaY96515
49	57	58.8	14	3	AaB16962
50	57	58.8	14	4	AaU25827
51	57	58.8	14	4	AaU26037
52	57	58.8	14	4	AaU26004
53	57	58.8	14	5	AbB72853
54	57	58.8	14	5	Abp51669
55	57	58.8	14	5	AaB18011
56	57	58.8	14	6	Abg71747
57	57	58.8	14	7	AbB62907
58	57	58.8	14	7	AdC33697
59	57	58.8	14	7	AdN59652
60	57	58.8	14	8	AdL27293
61	57	58.8	14	8	AdM72503
62	57	58.8	14	8	AdM72483
63	57	58.8	14	8	AdM72526
64	57	58.8	14	8	AdM72487
65	57	58.8	14	8	AdQ16584
66	57	58.8	15	2	AaW35416
67	57	58.8	15	2	AaW36780
68	57	58.8	15	2	AaW36776
69	57	58.8	15	2	AaW66714
70	57	58.8	15	2	AaW66721
71	57	58.8	15	2	AaW66712
72	57	58.8	15	3	AaB20684
73	57	58.8	15	4	AaU25996
74	57	58.8	15	4	AaU26026
75	57	58.8	15	4	AaU26020
76	57	58.8	15	4	AaU25831
77	57	58.8	15	4	AaU26007
78	57	58.8	15	5	Abp51670
79	57	58.8	15	7	AbB62908
80	57	58.8	15	8	AdM72485
81	57	58.8	15	8	AdM72479
82	57	58.8	15	8	AdM72502
83	57	58.8	15	8	AdM72478
84	57	58.8	15	8	AdM72492
85	57	58.8	15	8	AdM72533
86	57	58.8	15	8	AdM72490
87	57	58.8	15	8	AdM72486
88	57	58.8	15	8	AdM72491
89	57	58.8	15	8	AdM72523
90	57	58.8	15	8	AdM72493
91	57	58.8	15	8	AdM72482
92	57	58.8	15	8	AdQ16585
93	57	58.8	16	2	AaW19534
94	57	58.8	16	2	AaW33035
95	57	58.8	16	2	AaW36775
96	57	58.8	16	2	AaW36771
97	57	58.8	16	2	AaW36771
98	57	58.8	16	2	AaW66709

AaW09460	Thrombopo
AaW09498	Thrombopo
AaW36649	Thrombopo
AaW33027	Thrombopo
AaW36652	Thrombopo
AaB17026	TPO-mimet
AaU25868	Human thr
AaU25824	Human thr
AaU25871	Human thr
AbB72912	TPO mimet
AdJ73064	TPO mimet
AdJ52699	CHI delet
AdJ51660	CHI delet
AdQ16705	Modified
AdQ16704	Modified
AaW36779	Thrombopo
AaW09463	Thrombopo
AaW09468	Thrombopo
AaW33030	Thrombopo
AaW33034	Thrombopo
AaW36774	Thrombopo
AdJ24843	AP 12505
AaY96515	Thrombopo
AaB16962	TPO-mimet
AaU25827	Human thr
AaU26037	Human thr
AaU26004	Human thr
AbB72853	TPO mimet
Abp51669	Thrombopo
AaB18011	Human lig
Abg71747	TPO recep
AbB62907	Thrombopo
AdC33697	Erythro
AdN59652	Thrombopo
AdL27293	Amino aci
AdM72503	TPO mimet
AdM72483	TPO mimet
AdM72526	TPO mimet
AdM72487	TPO mimet
AdQ16584	Agonist T
AaW35416	Thrombopo
AaW36780	Thrombopo
AaW36776	Thrombopo
AaW66714	Peptide c
AaW66721	Peptide c
AaW66712	Peptide c
AbB20684	Thrombocy
AaU25996	Human thr
AaU26026	Human thr
AaU26020	Human thr
AaU25831	Human thr
AaU26007	Human thr
Abp51670	Thrombopo
AbB62908	Thrombopo
AdM72485	TPO mimet
AdM72479	TPO mimet
AdM72502	TPO mimet
AdM72478	TPO mimet
AdM72492	TPO mimet
AdM72533	TPO mimet
AdM72490	TPO mimet
AdM72486	TPO mimet
AdM72491	TPO mimet
AdM72523	TPO mimet
AdM72493	TPO mimet
AdM72482	TPO mimet
AdQ16585	TPO mimet
AaW19534	Thrombopo
AaW33035	Thrombopo
AaW36775	Thrombopo
AaW36771	Thrombopo
AaW66709	Peptide c

99 57 58.8 16 2 AAW66713 Peptide c  
100 57 58.8 16 2 AAW66733 Peptide c

## ALIGNMENTS

## RESULT 1

AAW09499 standard; protein; 18 AA.

XX AAW09499;  
AC AAW09499;  
XX 10-SEP-1997 (first entry)  
XX

Thrombopoietin receptor binding peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;  
KM bone marrow transfusion; chemotherapy; radiation therapy.  
XX

Synthetic.

W09640189-A1.

19-DEC-1996.

05-JUN-1996; 96WO-US008998.

07-JUN-1995; 95US-00472371.

07-JUN-1995; 95US-00473604.

07-JUN-1995; 95US-00476168.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00484090.

07-JUN-1995; 95US-00485301.

WPI; 1997-051883/05.

Thrombopoietin receptor-binding/activating peptide(s) and peptide

mimetic(s) - useful in treatment of haematological disorders, esp.

thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 27; 106pp; English.

The present sequence is a peptide which binds to thrombopoietin (TPO)

receptor (TR). The compound can be used for treating patients suffering

from haematological disorders and thrombocytopenia resulting from

chemotherapy, radiation therapy or bone marrow transfusions. The peptide

may also be used to maintain the proliferation and growth of TPO-

dependent cell lines and for use in biological research, for detecting

TPO receptors on living cells

Sequence 18 AA;

## RESULT 2

AAW09459 standard; protein; 18 AA.

AAW09459;

XX

10-SEP-1997 (first entry)

Thrombopoietin receptor binding compound peptide.

Haematology; thrombocytopenia; TPO; TR; proliferation;

bone marrow transfusion; chemotherapy; radiation therapy.

Synthetic.

Key Location/Qualifiers

Misc-difference 1..18 /note="Preferably linkages are selected from: -

CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6

; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is

lower alkyl"

Modified-site 1 /note="Preferably N-terminus is selected from: -NR1; -

NRC(O)R; -NRC(O)OR; -NRC(O)2R; -NHC(O)NR; succinimide;

benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3

substitutions on the phenyl ring selected from lower

alkyl, lower alkoxy, chloro, bromo; where R and R1 are

independently selected from hydrogen and lower alkyl"

Modified-site 18 /note="Preferably C-terminus is -C(O)R2 where R2 is

selected from hydroxy, lower alkoxy, and -NR3R4, where R3

and R4 are independently selected from hydrogen and lower

alkyl, and where the nitrogen atom of the -NR3R4 group

can optionally be the amine group of the N-terminus of

the peptide forming a cyclic peptide"

(GLAX ) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-051883/05.

Thrombopoietin receptor-binding/activating peptide(s) and peptide

mimetic(s) - useful in treatment of haematological disorders, esp.

thrombocytopenia resulting from chemotherapy, etc.

Claim 18; Page 89; 106pp; English.

The present sequence is a compound which binds to thrombopoietin (TPO)

receptor (TR). It has a molecular weight of &lt; 8000 Da, and a binding

affinity to TR as expressed by an IC50 of no more than about 100 nm. The

compound (especially if modified, see features table) can be used for

treating patients suffering from haematological disorders and

thrombocytopenia resulting from chemotherapy, radiation therapy or bone

marrow transfusions. The peptide may also be used to maintain the

proliferation and growth of TPO-dependent cell lines and for use in

biological research, for detecting TPO receptors on living cells

Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.7e-08; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TIKGPTLRQWLKSRHTS 18

Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 3  
AAW36650  
ID AAW36650 standard; peptide; 18 AA.

AC AAW36650;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PP 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Magstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.

PS Disclosure; Page 27; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

XX used to treat disorders which are susceptible to treatment with a

XX thrombopoietin agonist, preferably haematological disorders and

XX thrombocytopaenia resulting from chemotherapy, radiation therapy or bone

XX marrow transfusions. It can also be used diagnostically, e.g. to

XX investigate the mechanism of thrombopoietin signal transduction and

XX receptor activation, or to maintain the proliferation and growth of

XX thrombopoietin dependent cell lines

SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18

Db 1 TIKGPTLRQWLKSRHTS 18

KW Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PP 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Magstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.

PS Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a

XX molecular weight of less than 8000 Da and a TR binding affinity as

XX expressed by an IC50 of no more than about 100 microm, it can be used to

XX treat disorders which are susceptible to treatment with a thrombopoietin

XX agonist, preferably haematological disorders and thrombocytopaenia

XX resulting from chemotherapy, radiation therapy or bone marrow

XX transfusions. It can also be used diagnostically, e.g. to investigate the

XX mechanism of thrombopoietin signal transduction and receptor activation,

XX or to maintain the proliferation and growth of thrombopoietin dependent

XX cell lines

SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18

Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 5

AAW17024

ID AAW17024 standard; peptide; 18 AA.

AC AAW17024;

DT 31-OCT-2000 (first entry)

DE TPO-mimetic peptide sequence SEQ ID NO:80.

KW Modified peptide; therapeutic agent; fusion; Fe domain; cancer;

KW autoimmune disease; cystostatic; antiaesthatic; thrombolytic; VEGF;

KW immunosuppressive; BPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;

KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;

KW cytotoxic T cell lymphocyte antigen 4; tumor necrosis factor;

KW vascular endothelial growth factor; matrix metalloproteinase; asthma;

KW thrombosis; pharmaceutical.

OS Synthetic.

PN WO200024782-A2.

PD 04-MAY-2000.  
 XX  
 XX 25-OCT-1999; 99MO-US025044.  
 XX  
 PR 23-OCT-1998; 98US-0105371P.  
 XX 22-OCT-1999; 99US-00428082.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Feige U, Liu C, Cheetham J, Boone TC,  
 XX WPI, 2000-350702/30.  
 DR  
 XX  
 XX Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 XX Claim 19; Page 222; 608pp; English.  
 PS  
 XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-F1-(X2)-b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-C-P1, -(L1)-C-P1-(L2)-d-P2, -(L1)-C-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4, where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antisthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thromboestis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AA69443 to AA69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;  
 XX  
 Query Match 100.0%; Score 97; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRHTS 18  
 |||||  
 Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 6  
 AAU25869  
 ID AAU25869 standard; peptide; 18 AA.  
 XX  
 AC AAU25869;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #55.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX

PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96MO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX  
 XX Dower WJ, Barrett RM, Cwila SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagerstrom CR, Hendren RM, DePrince RM, Podduturi S,  
 PI Yin Q;  
 XX WPI, 2001-564142/63.  
 DR  
 XX  
 XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 XX Disclosure; Col 20; 128pp; English.  
 PS  
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX  
 SQ Sequence 18 AA;  
 XX  
 Query Match 100.0%; Score 97; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRHTS 18  
 |||||  
 Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 7  
 AAU25823  
 ID AAU25823 standard; peptide; 18 AA.  
 XX  
 AC AAU25823;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #9.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX

XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US0009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RM, Cwirla SE, Gates CM, Schatz PJ,  
 PI Balaubermanian P, Wagstrom CR, Hendren RM, Poddaturi S;  
 PI Yin Q;  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 67-68; 128pp; English.  
 XX  
 CC Sequences AU25815-AU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX  
 SQ Sequence 18 AA;  
 Query Match 100.0%; Score 97; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 TIKGPTLRQWLKSRHHTS 18  
 1 TIKGPTLRQWLKSRHHTS 18  
 Db 1 TIKGPTLRQWLKSRHHTS 18  
 RESULT 8  
 ABB72910  
 ID ABB72910 standard; peptide; 18 AA.  
 XX  
 AC ABB72910;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:80.  
 XX  
 KW Modified peptide; mimetic; Pc domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TWP;  
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antineumatic; antiarthritic; haemostatic; dermatological;  
 KW antianemic; anorectic; antinfertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX

OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200183525-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 02-MAY-2001; 2001WO-US014310.  
 XX  
 PR 03-MAY-2000; 2000US-00563286.  
 XX  
 PA (AMGEN-) AMGEN INC.  
 XX  
 PI Felge U, Liu C, Cheatham JC, Boone TC, Gudas JM;  
 PI WPI; 2002-130313/17.  
 DR  
 XX  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 44; 176pp; English.  
 XX  
 CC The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cyostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antinfertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-  
 CC infertility, and neurological degenerative diseases. (II), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;  
 Query Match 100.0%; Score 97; DB 5; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 TIKGPTLRQWLKSRHHTS 18  
 1 TIKGPTLRQWLKSRHHTS 18  
 Db 1 TIKGPTLRQWLKSRHHTS 18  
 RESULT 9  
 ADJ73062  
 ID ADJ73062 standard; peptide; 18 AA.  
 XX  
 AC ADJ73062;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE TPO mimetic peptide sequence Seqid 516.  
 XX  
 KW mimetic; CDR mimeticbody; gene therapy; transgenic; immune;  
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KW immunomodulator; cardiac; antimicrobial; cytostatic; neuroprotective;  
 KW TPO.  
 XX  
 OS Synthetic.

XX WO2003084477-A2.  
 XX 16-OCT-2003.  
 XX 24-MAR-2003; 2003WO-US009139.  
 XX 29-MAR-2002; 2002US-0368791P.  
 XX (CENZ ) CENTOCOR INC.  
 PA Heavner GA, Knight DM, Scallion BJ, Ghayeb J;  
 PI WPI; 2003-804237/75.  
 DR New CDR mimetibody comprising a portion of a heavy or light chain  
 XX variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PS cardiovascular or neurologic disease.  
 XX Disclosure; SEQ ID NO 516; 97pp; English.  
 CC This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LBR). The present invention describes human  
 CC CDR mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neuromodulator,  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 SO Sequence 18 AA;  
 Query Match 100.0%; Score 97; DB 7; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TIKGPTLRQWLKSRHNTS 18  
 DB 1 TIKGPTLRQWLKSRHNTS 18  
 RESULT 10  
 ID ADJ52697 standard; peptide; 18 AA.  
 AC ADJ52697;  
 XX 06-MAY-2004 (first entry)  
 DT CH1 deleted mimetibody-related peptide SeqID516.  
 XX CH1 deleted mimetibody-related peptide SeqID516.  
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arrhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX Unidentified.  
 OS Synthetic.  
 XX WO2004002417-A2.

PD 08-JUN-2004.  
 XX 27-JUN-2003; 2003WO-US020347.  
 XX 28-JUN-2002; 2002US-0392431P.  
 XX (CENZ ) CENTOCOR INC.  
 PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutooski KA;  
 DR WPI; 2004-082870/08.  
 XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 PS Claim 2; SEQ ID NO 516; 129pp; English.  
 CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 SO Sequence 18 AA;  
 Query Match 100.0%; Score 97; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TIKGPTLRQWLKSRHNTS 18  
 DB 1 TIKGPTLRQWLKSRHNTS 18  
 RESULT 11  
 ID ADJ51658 standard; peptide; 18 AA.  
 AC ADJ51658;  
 XX 06-MAY-2004 (first entry)  
 DT CH1 deleted mimetibody-related peptide SeqID516.  
 XX CH1 deleted mimetibody-related peptide SeqID516.  
 KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
 KW ophthalmological; nephroretropic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstetric disorder; hematologic disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 XX Unidentified.  
 OS

OS Synthetic.  
 XX  
 PN WO2004002424-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neseppor TC;  
 PI Kutolowski KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 15; SEQ ID NO 516; 123pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulatory, anti-allergic, muscular-Gen, cytostatic,  
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, allergic, infectious,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 XX  
 XX  
 SQ Sequence 18 AA;  
 Query Match 100.0%; Score 97; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TIKGPTLRQWLKSRHETS 18  
 DB 1 TIKGPTLRQWLKSRHETS 18  
 RESULT 12  
 AAW09491  
 ID AAW09491 standard; protein; 19 AA.  
 XX  
 AC AAW09491;  
 XX  
 DT 10-SEP-1997 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Haematology; thrombocytopenia; TPO; TR; proliferation;  
 KW bone marrow transfusion; chemotherapy; radiation therapy.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9640189-A1.  
 XX

PD 19-DEC-1996.  
 XX  
 PF 05-JUN-1996; 96WO-US008998.  
 XX  
 PR 07-JUN-1995; 95US-00472371.  
 PR 07-JUN-1995; 95US-00473604.  
 PR 07-JUN-1995; 95US-00476168.  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00484090.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX  
 DR WPI; 1997-051883/05.  
 XX  
 PT Thrombopoietin receptor-binding/activating peptide(s) and peptide  
 PT mimetic(s) - useful in treatment of haematological disorders, esp.  
 PT thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Disclosure; Page 26; 106pp; English.  
 XX  
 CC The present sequence is a peptide which binds to thrombopoietin (TPO)  
 CC receptor (TR). The compound can be used for treating patients suffering  
 CC from haematological disorders and thrombocytopenia resulting from  
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide  
 CC may also be used to maintain the proliferation and growth of TPO-  
 CC dependent cell lines and for use in biological research, for detecting  
 CC TPO receptors on living cells  
 XX  
 SQ Sequence 19 AA;  
 Query Match 64.9%; Score 63; DB 2; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0052;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GPTLRQWLKSRHETS 18  
 DB 5 GPTLRQWLKSRHETS 19  
 RESULT 13  
 AAW35418  
 ID AAW35418 standard; peptide; 19 AA.  
 XX  
 AC AAW35418;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Cross-links 3  
 FT /note= "linked via disulfide bond to Cys3 of identical  
 FT peptide"  
 FT Modified-site 19  
 FT /note= "NH2-Ser"  
 XX  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX

PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower MJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 9; Page 73; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 SQ Sequence 19 AA;  
 XX  
 Query Match 64.9%; Score 63; DB 2; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0052;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GPTLRQWLKSRHNTS 18  
 DB 5 GPTLRQWLARNHLS 19  
 XX  
 RESULT 14  
 AAW36642  
 ID AAW36642 standard; peptide; 19 AA.  
 XX  
 AC AAW36642;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 DE Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 OS Synthetic.  
 OS  
 XX  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower MJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX

PS Disclosure; Page 26; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 SQ Sequence 19 AA;  
 XX  
 Query Match 64.9%; Score 63; DB 2; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0052;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GPTLRQWLKSRHNTS 18  
 DB 5 GPTLRQWLARNHLS 19  
 XX  
 RESULT 15  
 AAU25861  
 ID AAU25861 standard; peptide; 19 AA.  
 XX  
 AC AAU25861;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #47.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescent-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower MJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;  
 PI Yan Q;  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hemtological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The



CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 19 AA;

Query Match 64.9%; Score 63; DB 4; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0052;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 GPTLRQWLKSRHNTS 18  
 |||||  
 DB 5 GPTLRQWLAAARNHLS 19

#### RESULT 16

AAU25998  
 ID AAU25998 standard; peptide; 19 AA.

AC AAU25998;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #184.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lact gene.

OS Homo sapiens.

PN US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Depince RB, Podduturi S;  
 PI Yin Q;

DR WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 143-144; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 19 AA;

Query Match 64.9%; Score 63; DB 4; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0052;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 GPTLRQWLKSRHNTS 18  
 |||||  
 DB 5 GPTLRQWLAAARNHLS 19

#### RESULT 17

ABP51693  
 ID ABP51693 standard; peptide; 18 AA.

AC ABP51693;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:49.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarily determining region; immunoglobulin; antianemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

PA (ALEXION) ALEXION PHARM INC.

PI Bowdish KS, Barbac-Frederickson S, Renshaw M;

DR WPI; 2002-566610/60.

DR N-PSDB; ABQ73371.

PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (1) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (1) has  
 CC antianemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX Sequence 18 AA:

Query Match 63.9%; Score 62; DB 5; Length 18;  
 Best Local Similarity 78.6%; Pred. No. 0.007;  
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14  
 2 TIKGPTLRQWLKSR 15

RESULT 18

ABP51691

ID ABP51691 standard; peptide; 18 AA.  
 AC ABP51691;  
 XX  
 DT 01-OCT-2002 (first entry)  
 XX

DE TPO mimetic peptide SEQ ID NO:45.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antinaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.  
 OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-028889P.

PR 29-MAY-2001; 2001US-0294068P.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Barbaa-Fredrickson S, Renshaw M;

XX WPI; 2002-566610/60.

DR N-PSDB; ABQ73369.

XX A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has

CC antinaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX Sequence 18 AA:

Query Match 63.9%; Score 62; DB 5; Length 18;  
 Best Local Similarity 78.6%; Pred. No. 0.007;  
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14  
 2 TIKGPTLRQWLKSR 15

RESULT 19

ADQ16625

ID ADQ16625 standard; peptide; 18 AA.  
 AC ADQ16625;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX

DE TPO mimetic peptide with random flanking residues SEQ ID NO:45.

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.

OS Unidentified.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;

XX WPI; 2004-460973/43.

DR N-PSDB; ADQ16626.

XX New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX Example 1; SEQ ID NO 45; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow

CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide with flanking  
CC residues.  
XX  
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 8; Length 18;  
Best Local Similarity 78.6%; Pred. No. 0.007;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSR 14  
||:|||||:|  
Db 2 TIEGPTLRQWLAAAR 15

RESULT 20  
ADQ16629  
ID ADQ16629 standard; peptide; 18 AA.  
XX  
AC ADQ16629;  
XX  
DT 09-SEP-2004 (first entry)  
XX

DE TPO mimetic peptide with random flanking residues SEQ ID NO:49.

XX immunoglobulin; complementarily determining region; CDR; peptide mimetic;  
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
KW immunotherapy; thrombocytopenia.

XX Unidentified.

OS  
PN WO2004050017-A2.

XX 17-JUN-2004.

XX 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;

XX WPI; 2004-460973/43.

DR N-PSDB; ADQ16630.

XX New immunoglobulin molecule comprising a region, where two  
PT complementarily determining regions (CDRs) are replaced with EPO mimetic  
PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 1; SEQ ID NO 49; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
CC comprising a region where amino acid residues corresponding to at least a  
CC portion of a two complementarily determining regions (CDRs) are replaced  
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
CC invention has immunosuppressive activity, and may have a use in  
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide with flanking  
CC residues.

SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 8; Length 18;  
Best Local Similarity 78.6%; Pred. No. 0.007;  
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSR 14  
||:|||||:|  
Db 2 TIEGPTLRQWLAAAR 15

## RESULT 21

AAW09493  
ID AAW09493 standard; protein; 19 AA.

XX  
AC AAW09493;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;  
KW bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX WO9640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

XX 07-JUN-1995; 95US-00473604.

XX 07-JUN-1995; 95US-00476168.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00484090.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirja SE, Duffin DJ, Gates CM, Johnson SS;

XX Matcheakts LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

XX mimetic(s) - useful in treatment of haematological disorders, esp.

XX thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 26; 106pp; English.

XX Sequence 19 AA;

Query Match 63.9%; Score 62; DB 2; Length 19;  
Best Local Similarity 73.3%; Pred. No. 0.0075;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 GPTLRQWLKSRHNTS 18  
|||||||:|  
Db 5 GPTLRQWLAAARTHLS 19

## RESULT 22

AAW36644  
ID AAW36644 standard; peptide; 19 AA.

XX  
AC AAW36644;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 XX WO9640750-A1.  
 PN  
 XX 19-DEC-1996.  
 PD  
 XX  
 XX 07-JUN-1996; 96WO-US009623.  
 PF  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PA  
 XX (GLAXO ) GLAXO GROUP LTD.  
 PI Dower MJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Metheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 DR WPI; 1997-052226/05.  
 XX  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 XX Disclosure; Page 26; 106pp; English.  
 PS  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 19 AA;  
 XX

Query Match 63.9%; Score 62; DB 2; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0075;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRHTS 18  
 ||||| : |||  
 Db 5 GPTLRQWLARTHS 19

RESULT 23  
 AAU25863  
 ID AAU25863 standard; peptide; 19 AA.  
 XX  
 AC AAU25863;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #49.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 XX 26-JUN-2001.  
 PD  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX

PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 XX (GLAXO ) GLAXO GROUP LTD.  
 PA  
 XX Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balaubramanian P, Wagstrom CR, Hendren RM, Deprience RB, Podduuri S;  
 PI Yin Q;  
 XX  
 XX WPI; 2001-564142/63.  
 DR  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprising contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 XX Disclosure; Col 20; 128pp; English.  
 PS  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX  
 SQ Sequence 19 AA;  
 XX

Query Match 63.9%; Score 62; DB 4; Length 19;  
 Best Local Similarity 73.3%; Pred. No. 0.0075;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRHTS 18  
 ||||| : |||  
 Db 5 GPTLRQWLARTHS 19

RESULT 24  
 ABG71748  
 ID ABG71748 standard; protein; 144 AA.  
 XX  
 XX ABG71748;  
 AC  
 XX  
 DT 20-JAN-2003 (first entry)  
 XX  
 DE Antibody CDR containing MPL-TPO binding sequence, TPOVHCDR1.  
 XX  
 KW Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;  
 KW complementarily determining region; CDR; antigenic; thrombopoietin; TPO;  
 KW thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;  
 KW T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;  
 KW proliferation; growth; differentiation; haematopoietic cell; antibody;  
 KW platelet progenitor cell; immune disorder; thrombocytopenia;  
 KW disseminated intravascular coagulation; stem cell; transplantation;  
 KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant;  
 KW consensus variable heavy chain domain; CONVH.  
 XX  
 OS Synthetic.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH

FT Region 50..63  
 XX /note= "TPO receptor binding agonist peptide"  
 PN WO200278612-A2.  
 PD 10-Oct-2002.  
 PR 02-Apr-2002; 2002WO-US010301.  
 PR 02-Apr-2001; 2001US-0281183P.  
 XX (PURD ) PURDUE PHARMA LP.  
 PA Solcis DA, Burch RM, Ogert RA;  
 PI WPI; 2003-040615/03.  
 DR  
 XX New thrombopoietin synthebodies, useful for stimulating proliferation,  
 PT growth, or differentiation of hematopoietic cells, for treating or  
 PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.  
 PS Example 1; Page 75; 97pp; English.

The invention discloses a variant of an immunoglobulin (Ig) variable heavy or light chain domain that comprises at least one complementary determining region (CDR) and framework regions flanking the CDR. The CDR also has added or substituted to it, at least one binding sequence which is heterologous to the CDR and is an antigenic, agonistic sequence from a thrombopoietin (TPO) receptor binding sequence. The antigenic sequence can be a binding sequence heterologous to the CDR, a cytotoxic T-lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper cell sequence or a combination of each. The variant or thrombopoietin synthebody, pharmaceutical and vaccine compositions are useful for stimulating proliferation, growth or differentiation of haematopoietic cells; particularly platelet progenitor cells. The variants are also useful for treating or preventing haematopoietic or immune disorders resulting from chemotherapy, radiation therapy, or bone marrow transfusions (e.g. thrombocytopenia or disseminated intravascular coagulation). Compositions comprising the synthebodies can be used for the mobilisation, amplification and ex vivo expansion of stem cells and committed precursor cells for autologous and allogeneic transplantation as well as for the expansion of stem cells for gene therapy. They are also useful as diagnostic or analytical reagents for studying the function of thrombopoietin and its receptor in vivo or in vitro. The sequence presented, TPVHCRL, is the TPO receptor (MPJ) agonist peptide sequence contained within the immunoglobulin CDR consensus variable heavy chain domain (CONVH)

Sequence 144 AA:  
 SQ

Query Match	Score 62;	DB 6;	Length 144;
Best Local Similarity	78.6%	Pred. No. 0.076;	
Matches 11; Conservative	2;	Mismatches 1;	Indels 0; Gaps 0

```
Qy      1  TIKGPTLRQLWLSKR 14
       ||:|||||:|
Db     49  TIEGTLRQLWLAR 62
```

RESULT 25  
 ID AUAZ6006  
 AC AUAZ6006; standard; peptide; 14 AA.  
 DT 17-DEC-2001 (first entry)  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #192.  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantion; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;

KM	tissue homogenate; fluorescence-activated cell sorting; Western blotting.
KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX	
OS	Homo sapiens.
PN	US6251864-B1.
PD	26-JUN-2001.
XX	
PF	01-MAR-2000; 2000US-00516704.
XX	
PR	07-JUN-1995; 95US-00478128.
PR	07-JUN-1995; 95US-00485301.
PR	07-JUN-1996; 96WO-US009623.
PR	15-AUG-1996; 96US-0069027.
XX	
PA	(GLAXO ) GLAXO GROUP LTD.
P1	Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schartz PJ,
P1	Balasubdramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;
P1	Yin Q;
DR	WPI; 2001-564142/63.
XX	
PT	Activating thrombopoietin receptors in cells, used to treat
PT	thrombocytopenia and hematological disorders, comprises contacting cells
PT	with peptides and peptide mimetics attached to hydrophilic polymers.
XX	
PS	Disclosure; Col 147; 128pp; English.
XX	
CC	Sequences AAN25815-AAU26049 represent peptides and peptide mimetics that
CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC	of activating thrombopoietin receptors in cells comprise contacting the
CC	cells with effective amounts of peptides and peptide mimetics attached to
CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC	as that due to chemotherapy, radiation therapy or bone-marrow
CC	transplantation and to prevent thrombocytopenia in patients at risk. The
CC	sequences are used to treat and prevent haematological disorders
CC	including thrombocytopenia and platelet disorders. They are used in vitro
CC	as unique tools for understanding the biological role of thrombopoietin
CC	(TPO) and to develop other compounds that bind to and activate the TPO
CC	receptor. The peptides can be used to detect TPO receptors on living
CC	cells and fixed cells, in biological fluids, in tissue homogenates, and
CC	in purified or natural biological materials. They may also be used for in
CC	situ staining, fluorescence-activated cell sorting, Western blotting and
CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC	be used for in vitro expansion of megakaryocytes and their committed
CC	progenitors alone or in conjunction with additional cytokines
XX	
SQ	Sequence 14 AA;
XX	
Query Match	59.8%; Score 58; DB 4; Length 14;
Best Local Similarity	71.4%; Pred. No. 0.023; Mismatches 1; Indels 0; Gaps 0;
Matches	10; Conservative 3;
Oy	2 IKGPTLRQWLKRSRE 15  ::   ::  Db 1 IEGPTLRQWLAKRK 14
RESULT 26	
ID AAM09460	AAM09460 standard; protein, 18 AA.
AC AAM09460;	
XX	
DT 10-SEP-1997	(first entry)
XX	
DE Thrombopoietin receptor binding compound peptide.	
XX	
KM Haematology; thrombocytopenia; TPO; TR; proliferation;	
XX	bone marrow transfusion; Chemotherapy; radiation therapy.
XX	

```

OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 1.18
FT /note= "Preferably linkages are selected from: -
FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
FT lower alkyl"
FT Modified-site
FT 1
FT /note= "Preferably N-terminus is selected from: -NRR1; -
FT NRC(O)R; -NRC(O)R; -NRC(O)2R; -NRC(O)NR; succinimide;
FT benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3
FT substitutions on the phenyl ring selected from lower
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
FT independently selected from hydrogen and lower alkyl"
FT Modified-site
FT 18
FT /note= "Preferably C-terminus is -C(O)R2 where R2 is
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
FT and R4 are independently selected from hydrogen and lower
FT alkyl, and where the nitrogen atom of the -NR3R4 group
FT can optionally be the amine group of the N-terminus of
FT the peptide forming a cyclic peptide"
XX
XX W09640189-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US008998.
XX
XX 07-JUN-1995; 95US-00472371.
XX 07-JUN-1995; 95US-00473604.
XX 07-JUN-1995; 95US-00476168.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00484090.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 18; Page 89; 106pp; English.
XX
XX The present sequence is a compound which binds to thrombopoietin (TPO)
XX receptor (TR). It has a molecular weight of < 8000 Da, and a binding
XX affinity to TR as expressed by an IC50 of no more than about 100 nm. The
XX compound (especially if modified, see features table) can be used for
XX treating patients suffering from haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. The peptide may also be used to maintain the
XX proliferation and growth of TPO-dependent cell lines and for use in
XX biological research, for detecting TPO receptors on living cells
XX
XX Sequence 18 AA;
XX
XX Query Match 59.8%; Score 58; DB 2; Length 18;
XX Best Local Similarity 71.4%; Pred. No. 0.031;
XX Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TIKPTLRQWLKSR 14
XX :|:|||||:|
XX 1 SIEGPTLRWLTSR 14
XX
XX RESULT 27
XX AAW09498
XX ID AAW09498 standard; protein; 18 AA.

```

```

XX
XX AAW09498;
XX
XX 10-SEP-1997 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.
XX
XX Synthetic.
XX
XX W09640189-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US008998.
XX
XX 07-JUN-1995; 95US-00472371.
XX 07-JUN-1995; 95US-00473604.
XX 07-JUN-1995; 95US-00476168.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00484090.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 27; 106pp; English.
XX
XX The present sequence is a peptide which binds to thrombopoietin (TPO)
XX receptor (TR). The compound can be used for treating patients suffering
XX from haematological disorders and thrombocytopenia resulting from
XX chemotherapy, radiation therapy or bone marrow transfusions. The peptide
XX may also be used to maintain the proliferation and growth of TPO-
XX dependent cell lines and for use in biological research, for detecting
XX TPO receptors on living cells
XX
XX Sequence 18 AA;
XX
XX Query Match 59.8%; Score 58; DB 2; Length 18;
XX Best Local Similarity 71.4%; Pred. No. 0.031;
XX Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TIKPTLRQWLKSR 14
XX :|:|||||:|
XX 1 SIEGPTLRWLTSR 14
XX
XX RESULT 28
XX AAW36649
XX ID AAW36649 standard; peptide; 18 AA.
XX
XX AAW36649;
XX
XX 11-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX

```

```

XX PN W09640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower MJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX PT Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 27; 106pp; English.
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX CC used to treat disorders which are susceptible to treatment with a
XX CC thrombopoietin agonist, preferably haematological disorders and
XX CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX CC marrow transfusions. It can also be used diagnostically, e.g. to
XX CC investigate the mechanism of thrombopoietin signal transduction and
XX CC receptor activation, or to maintain the proliferation and growth of
XX CC thrombopoietin dependent cell lines
XX SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.031;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
   :|:|||||:|
Db 1 SIEGPTLRWLTSLR 14

RESULT 29
AAW33027
ID AAW33027 standard; peptide; 18 AA.
XX
XX AC AAW33027;
XX DT 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
XX KW haematological disorder; thrombocytopenia; chemotherapy;
XX KW radiation therapy; bone marrow transfusion; diagnosis;
XX KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX PN W09640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower MJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

```

```

XX DR WPI; 1997-052226/05.
XX PT Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Claim 19; Page 89; 106pp; English.
XX CC The present peptide binds the thrombopoietin receptor (TR), has a
XX CC molecular weight of less than 8000 Da and a TR binding affinity as
XX CC expressed by an IC50 of no more than about 100 microm. It can be used to
XX CC treat disorders which are susceptible to treatment with a thrombopoietin
XX CC agonist, preferably haematological disorders and thrombocytopenia
XX CC resulting from chemotherapy, radiation therapy or bone marrow
XX CC transfusions. It can also be used diagnostically, e.g. to investigate the
XX CC mechanism of thrombopoietin signal transduction and receptor activation,
XX CC or to maintain the proliferation and growth of thrombopoietin dependent
XX CC cell lines
XX SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.031;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
   :|:|||||:|
Db 1 SIEGPTLRWLTSLR 14

RESULT 30
AAW3652
ID AAW3652 standard; peptide; 18 AA.
XX
XX AC AAW3652;
XX DT 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
XX KW haematological disorder; thrombocytopenia; chemotherapy;
XX KW radiation therapy; bone marrow transfusion; diagnosis;
XX KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX PN W09640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower MJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX PT Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 27; 106pp; English.
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX CC used to treat disorders which are susceptible to treatment with a
XX CC thrombopoietin agonist, preferably haematological disorders and

```



CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.031;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQMLKSR 14  
 :|||:|  
 Db 1 SIEGPTLRRLWTSR 14

RESULT 31

AAB17026  
 ID AAB17026 standard; peptide; 18 AA.

AC AAB17026;

DT 31-OCT-2000 (first entry)

DE TPO-mimetic peptide sequence SEQ ID NO:82.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytotoxic; antineoplastic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.

OS Synthetic.

PN WO200024782-A2.

PD 04-MAY-2000.

PF 25-OCT-1999; 99WO-US025044.

PR 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

PA (AMGE-) AMGEN INC.

PI Feige U, Liu C, Cheetham J, Boone TC;

DR WPI; 2000-350702/30.

PT Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.

PS Claim 19; Page 222; 608pp; English.

XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)d-P2, -(L1)-c-P1-  
 CC (L2)d-P2-(L3)e-P3, or -(L1)-c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antineoplastic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to

CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention

XX Sequence 18 AA;

Query Match 59.8%; Score 58; DB 3; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.031;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQMLKSR 14  
 :|||:|  
 Db 1 SIEGPTLRRLWTSR 14

RESULT 32

AAU25868  
 ID AAU25868 standard; peptide; 18 AA.

AC AAU25868;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #54.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS Homo sapiens.

PN US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;

PI Balaubramanian P, Wagstrom CR, Hendren RW, Deprience RB, Poddaturi S;

PI yin Q;

DR WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and



CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 4; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.031;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 TIKGPTLRQWLKSR 14  
 :|:|||||:|  
 Db 1 SIEGPTLRQWLKSR 14

## RESULT 33

AAU25824  
 ID AAU25824 standard; peptide; 18 AA.

AC AAU25824;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #10.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

PI Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Poddaturi S;

PI Yln Q;

DR WPI; 2001-564142/63.

PS Disclosure; Col 67-68; 128pp; English.

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 4; Length 18;  
 Best Local Similarity 71.4%; Pred. No. 0.031;  
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 TIKGPTLRQWLKSR 14  
 :|:|||||:|  
 Db 1 SIEGPTLRQWLKSR 14

## RESULT 34

AAU25871  
 ID AAU25871 standard; peptide; 18 AA.

AC AAU25871;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #57.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

PI Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Poddaturi S;

PI Yln Q;

DR WPI; 2001-564142/63.

PS Disclosure; Col 20; 128pp; English.

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in  
CC situ staining, fluorescence-activated cell sorting, western blotting and  
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
CC be used for in vitro expansion of megakaryocytes and their committed  
CC progenitors alone or in conjunction with additional cytokines  
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 4; Length 18;  
Best Local Similarity 71.4%; Pred. No. 0.031;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQMLKSR 14  
:|:|||||:|  
Db 1 SIEGPTLRWLTSR 14

RESULT 35  
ID ABB72912 standard; peptide; 18 AA.  
XX ABB72912;

AC ABB72912;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:82.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
KM erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
KM TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
KM TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;  
KM MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
KM cyclostatic; antirheumatic; antidiabetic; ophthalmological;  
KM antinaeemic; anorectic; antifertility; haemostatic; dermatological;  
KM neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
KM cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
KM sleep disorder; neurological degenerative disease; anaemia;  
KM thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
KM Fanconi's syndrome.

OS Homo sapiens.  
OS Synthetic.

PN WO200183525-A2.

PD 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGEN-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

PT Novel vehicle-peptide molecule or its multimers useful for treating  
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 44; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its  
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
CC cyostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
CC antinaeamic, anorectic, antifertility, haemostatic, dermatological and  
CC neuroprotective activities. (I) can be used as a therapeutic or  
CC prophylactic agent as well as for screening purposes. (I) is useful for  
CC diagnosing diseases characterised by dysfunction of their associated  
CC protein of interest, for identifying normal or abnormal proteins of  
CC interest, as a part of diagnostic kit to detect the presence of their  
CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, cancer,  
CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
CC mimetic compounds are useful for treating disorders characterised by low  
CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
CC compounds are useful for treating conditions that involve an existing  
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABB72403 to ABB73426 and AB135695 to AB135777  
CC represent amino acid and nucleic acid sequences used in the  
CC exemplification of the present invention

SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 5; Length 18;  
Best Local Similarity 71.4%; Pred. No. 0.031;  
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQMLKSR 14  
:|:|||||:|  
Db 1 SIEGPTLRWLTSR 14

RESULT 36  
ADJ73064 standard; peptide; 18 AA.  
XX ADJ73064;

AC ADJ73064;

DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 518.

XX mimetic; CDR mimeticbody; gene therapy; transgenic; immune;  
KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
KM immunomodulator; cardiant; antimicrobial; cyostatic; neuroprotective;  
KM TPO.

OS Synthetic.

PN WO2003084477-A2.

PD 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENTZ ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

PT New CDR mimeticbody comprising a portion of a heavy or light chain  
PT variable region comprising human framework or ligand binding region,  
PT useful for preparing a composition for treating e.g., immune,  
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 518; 97pp; English.

XX This invention relates to novel mammalian CDR mimeticbodies, specific  
CC portions or variants thereof. Specifically, it refers to an antibody  
CC fragment where a protein has been inserted into, or replaces a portion  
CC of, one or more CDR regions, such that each CDR mimeticbody comprises at  
CC least one portion of a heavy chain or light chain variable region, which  
CC itself comprises at least one human framework region and at least one  
CC ligand binding region (LBR). The present invention describes human  
CC mimeticbodies, including modified immunoglobulins and cleavage products  
CC that can be useful in gene therapy and the generation of transgenic  
CC plants and animals. Furthermore, the CDR mimeticbody is useful for  
CC preparing compositions for modulating, treating or reducing the symptoms



CC	modulator or cytokine-agonist. The methods and compositions of the
CC	present invention are useful for the diagnosis, prevention and/or
CC	treatment of diseases or conditions associated with aberrant expression
CC	or activity of the CH1 deleted mimetobody, such as a bone or joint,
CC	cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC	endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC	obstetric, haematologic, immunological, allergic, infectious,
CC	musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC	pediatric, psychiatric, renal or pulmonary disorders. The present
CC	sequence is that of a peptide which may be used during the creation of a
CC	mimetbody of the invention.
XX	
SQ	Sequence 18 AA;
QY	Query Match                 59.8%;      Score 58;      DB 8;      Length 18;
	Best Local Similarity       71.4%;      Pred. No. 0.031;
Matches	10; Conservative          3; Mismatches    1; Indels      0; Gaps      0;
Dd	1 TIKGPTLRQWLKSR 14 :::     ::   1 SIEGPTLRWLTSLR 14
RESULT 39	
ADQ16705	
ID	ADQ16705 standard; protein; 128 AA.
XX	
AC	ADQ16705;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Modified immunoglobulin clone 116 HC variable region SEQ ID NO:125.
KM	immunoglobulin; complementarity determining region; CDR; peptide mimetic;
XV	erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW	immunotherapy; thrombocytopenia.
XX	
OS	Synthetic.
XX	
PN	WO2004050017-A2.
PD	17-JUN-2004.
XX	
PE	17-NOV-2003; 2003MO-US036894.
XX	
FR	02-DEC-2002; 2002US-00307724.
XX	
PA	(ALEX-) ALEXION PHARM INC.
PI	Bowdish KS, Frederickson S, Renshaw M;
XX	
DR	WPI; 2004-460973/43.
XX	
PT	New immunoglobulin molecule comprising a region, where two
PT	complementarity determining regions (CDRs) are replaced with EPO mimetic
PT	or a TPO mimetic, useful for treating thrombocytopenia.
XX	
Claim 9;	SEQ ID NO 125; 107bp; English.
PS	
XX	
CC	The invention relates to a novel immunoglobulin molecule or its fragment
CC	comprising a region where amino acid residues corresponding to at least a
CC	portion of a two complementarity determining regions (CDRs) are replaced
CC	with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC	a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC	invention has immunosuppressive activity, and may have a use in
CC	immunochemistry. The immunoglobulin molecule is useful for diagnosing or
CC	treating thrombocytopenia as a result of chemotherapy, bone marrow
CC	transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC	The present sequence represents immunoglobulin clone 116 heavy chain
XX	variable region.
XX	
SQ	Sequence 128 AA;

```

Query Match          59.8%; Score 58; DB 8; Length 128;
Best Local Similarity 62.5%; Pred. No. 0.29;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2  IKGPTLRQWLKSRHEHT 17
      |:|||||||:|::
Db      52  IEGPTLRQWLARANS 67

RESULT 40
ADQ16704
ID      ADQ16704 standard; protein; 225 AA.
AC
ADQ16704;
DT      09-SEP-2004 (first entry)
DE      Modified immunoglobulin clone 116 heavy chain SEQ ID NO:124.
KW      immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW      erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW      immunotherapy; thrombocytopenia.
XX      Synthetic.
OS
XX      WO2004050017-A2.
XX      17-JUN-2004.
XX      17-NOV-2003; 2003WO-US036894.
XX      02-DEC-2002; 2002US-00307724.
XX      (ALEX-) ALEXION PHARM INC.
XX      Bowdish KS, Frederickson S, Renshaw M;
XX      WPI; 2004-460973/43.
XX      New immunoglobulin molecule comprising a region, where two
XX      complementarity determining regions (CDRs) are replaced with EPO mimetic
XX      or a TPO mimetic, useful for treating thrombocytopenia.
XX      Example 8; SEQ ID NO 124; 107bp; English.
XX      The invention relates to a novel immunoglobulin molecule or its fragment
XX      comprising a region where amino acid residues corresponding to at least a
XX      portion of a two complementarity determining regions (CDRs) are replaced
XX      with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
XX      a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
XX      invention has immunosuppressive activity, and may have a use in
XX      immunotherapy. The immunoglobulin molecule is useful for diagnosing or
XX      treating thrombocytopenia as a result of chemotherapy, bone marrow
XX      transplantation, or chronic diseases such as idiopathic thrombocytopenia.
XX      The present sequence represents immunoglobulin clone 116 heavy chain.
XX      Sequence 225 AA;

Query Match          59.8%; Score 58; DB 8; Length 225;
Best Local Similarity 62.5%; Pred. No. 0.56;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2  IKGPTLRQWLKSRHEHT 17
      |:|||||||:|::
Db      52  IEGPTLRQWLARANS 67

RESULT 41
AAM36779
ID      AAM36779 standard; peptide; 13 AA.
AC
AAM36779;
XX

```

DT		11-MAR-1998 (first entry)	
XX			
DE		Thrombopoietin receptor binding peptide.	
XX			
KW		Thrombopoietin receptor; binding peptide; treatment; agonist;	
KV		haematological disorder; thrombocytopaenia; chemotherapy;	
KW		radiation therapy; bone marrow transfusion; diagnosis;	
KM		signal transduction; receptor activation; cell culture.	
XX			
OS		Synthetic.	
XX			
FH	Key	Location/Qualifiers	
FT	Cross-link	13	
FT		/note= "terminal carboxy group linked to epsilon amino	
FT		group of Lys15 in AAM36780"	
XX			
PN	MO9640750-A1.		
PD	19-DEC-1996.		
XX			
PF	07-JUN-1996;	96WO-US009623.	
XX			
PR	07-JUN-1995;	95US-00478128.	
XX			
PR	07-JUN-1995;	95US-00485301.	
XX			
PA	(GLAX ) GLAXO GROUP LTD.		
XX			
P1	Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;		
PI	Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;		
XX			
DR	WPI, 1997-052226/05.		
XX			
PT	Peptides and peptide mimetics which bind to and activate the		
PT	thrombopoietin receptor - useful in treatment of haematological		
PT	disorders, esp. thrombocytopenia resulting from chemotherapy, etc.		
XX			
PS	Example 9; Page 78; 106pp; English.		
XX			
CC	The present peptide, which binds the thrombopoietin receptor (TR), can be		
CC	used to treat disorders which are susceptible to treatment with a		
CC	thrombopoietin agonist, preferably haematological disorders and		
CC	thrombocytopenia resulting from chemotherapy, radiation therapy or bone		
CC	marrow transfusions. It can also be used diagnostically, e.g. to		
CC	investigate the mechanism of thrombopoietin signal transduction and		
CC	receptor activation, or to maintain the proliferation and growth of		
CC	thrombopoietin dependent cell lines		
XX			
SQ	Sequence 13 AA:		
	Query Match	58.8%; Score 57; DB 2; Length 13;	
	Best Local Similarity	76.9%; Pred. No. 0.03;	
	Matches 10; Conservative	2; Mismatches 1; Indels 0; Gaps 0.	
DY			
	2 IKGPTLRQWLKSR 14		
	:               :		
	1 IEGPLRQLAAR 13		
ID			
AAW09463	standard; protein; 14 AA.		
XX			
AC	AAW09463;		
XX			
DT			
XX	10-SEP-1997 (first entry)		
DE			
XX	Thrombopoietin receptor binding compound peptide.		
XX			
KW	Haematology; thrombocytopenia; TPO; TR; proliferation;		
KW	bone marrow transfusion; chemotherapy; radiation therapy.		
XX			
OS	Synthetic.		
XX			

Key	Location/Qualifiers
FT	1. 14
FT	/note= "Preferably linkages are selected from: -
FT	CH2C(O)NR-; phosphonate; -CH2S(O)NR-; -CH2NR-; -C(O)NR6
FT	; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
FT	lower alkyl"
FT	1
FT	/note= "Preferably N-terminus is selected from: -NRR1; -
FT	NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
FT	benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3
FT	substitutions on the phenyl ring selected from lower
FT	alkyl, lower alkoxy, chloro, bromo; where R and R1 are
FT	independently selected from hydrogen and lower alkyl"
FT	14
FT	/note= "Preferably C-terminus is -C(O)R2 where R2 is
FT	selected from hydroxy, lower alkoxy, and -NR3R4, where R3
FT	and R4 are independently selected from hydrogen and lower
FT	alkyl, and where the nitrogen atom of the -NR3R4 group
FT	can optionally be the amine group of the N-terminus of
FT	the peptide forming a cyclic peptide"
XX	
XX	WO9640189-A1.
XX	
XX	19-DEC-1996.
XX	
XX	05-JUN-1996; 96WO-US008998.
XX	
XX	07-JUN-1995; 95US-00472371.
XX	07-JUN-1995; 95US-00473604.
XX	07-JUN-1995; 95US-00476168.
XX	07-JUN-1995; 95US-00478128.
XX	07-JUN-1995; 95US-00484090.
XX	07-JUN-1995; 95US-00485301.
XX	
XX	(GLAXO ) GLAXO GROUP LTD.
XX	
XX	Dower MJ, Barrett RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX	Mattheakis LC, Schatz PJ, Wagerstrom CR, Wrighton NC;
XX	WPI; 1997-051883/05.
XX	
XX	Thrombopoietin receptor-binding/activating peptide (s) and peptide
XX	mimetic(s) - useful in treatment of haematological disorders, esp.
XX	thrombocytopenia resulting from chemotherapy, etc.
XX	
XX	Claim 18; Page 89; 106pp; English.
XX	
XX	The present sequence is a compound which binds to thrombopoietin (TPO)
XX	receptor (TR). It has a molecular weight of < 8000 Da, and a binding
XX	affinity to TR as expressed by an IC50 of no more than about 100 nM. The
XX	compound (especially if modified, see features table) can be used for
XX	treating patients suffering from haematological disorders and
XX	thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX	marrow transfusions. The peptide may also be used to maintain the
XX	proliferation and growth of TPO-dependent cell lines and for use in
XX	biological research, for detecting TPO receptors on living cells
XX	
XX	Sequence 14 AA;
XX	
XX	Query Match 58.8%; Score 57; DB 2; Length 14;
XX	Best Local Similarity 76.9%; Pred. No. 0.033; Mismatches 0; Gaps 0;
XX	Matches 10; Conservative 2; Indels 1;
XX	
XX	2 IKGPTLRQWLKSR 14
XX	1 IEGPTLRQWLAR 13
XX	
XX	RESULT 43
XX	AAW09468
XX	AAW09468 standard; protein; 14 AA.
XX	AAW09468;

```

XX 10-SEP-1997 (first entry)
XX Thrombopoietin receptor binding compound peptide (part of a dimer).
DE Thrombopoietin receptor binding compound peptide (part of a dimer).
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.
XX Synthetic.
XX Key Location/Qualifiers
FT Cross-links 14 /note= "linked to the omega lys from AAM19534"
XX
XX MO9640189-A1.
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US008998.
XX
XX 07-JUN-1995; 95US-00472371.
XX 07-JUN-1995; 95US-00473604.
XX 07-JUN-1995; 95US-00476168.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00484090.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 30; Page 91; 106pp; English.
XX
XX The present sequence is a compound which binds to thrombopoietin (TPO)
XX receptor (TR). It is part of a dimer linked by the omega amino acid to
XX the omega amino acid in the sequence in AAM19534. The compound can be
XX used for treating patients suffering from haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. The peptide may also be used to maintain the
XX proliferation and growth of TPO-dependent cell lines and for use in
XX biological research, for detecting TPO receptors on living cells
XX
XX Sequence 14 AA;
SQ
XX
XX Query Match 58.8%; Score 57; DB 2; Length 14;
XX Best Local Similarity 76.9%; Pred. No. 0.033;
XX Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 IKGPTLRQWLKSR 14
XX |:|||||:|
XX 1 IEKPTLRQWLKSR 13
DB
XX
XX RESULT 44
XX AAM33030
XX ID AAM33030 standard; peptide; 14 AA.
XX
XX AAM33030;
XX
XX 11-MAR-1998 (first entry)
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX
XX

```

```

XX signal transduction; receptor activation; cell culture.
XX Synthetic.
XX MO9640750-A1.
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 19; Page 89; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
XX molecular weight of less than 8000 Da and a TR binding affinity as
XX expressed by an IC50 of no more than about 100 microm. It can be used to
XX treat disorders which are susceptible to treatment with a thrombopoietin
XX agonist, preferably haematological disorders and thrombocytopenia
XX resulting from chemotherapy, radiation therapy or bone marrow
XX transfusions. It can also be used diagnostically, e.g. to investigate the
XX mechanism of thrombopoietin signal transduction and receptor activation,
XX or to maintain the proliferation and growth of thrombopoietin dependent
XX cell lines
XX
XX Sequence 14 AA;
SQ
XX
XX Query Match 58.8%; Score 57; DB 2; Length 14;
XX Best Local Similarity 76.9%; Pred. No. 0.033;
XX Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 IKGPTLRQWLKSR 14
XX |:|||||:|
XX 1 IEKPTLRQWLKSR 13
DB
XX
XX RESULT 45
XX AAM33034
XX ID AAM33034 standard; peptide; 14 AA.
XX
XX AAM33034;
XX
XX 11-MAR-1998 (first entry)
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Cross-links 14 /note= "terminal carboxy group linked to epsilon amino
XX group of Lys16 in AAM33035"
XX
XX MO9640750-A1.
XX 19-DEC-1996.
XX

```

```

XX 07-JUN-1996; 96WO-US009623.
PF
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAXO ) GLAXO GROUP LTD.
PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of hematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX
PS Claim 30; Page 91; 106pp; English.
XX
CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably hematological disorders and thrombocytopaenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 14 AA;

```

```

Query Match 58.8%; Score 57; DB 2; Length 14;
Best Local Similarity 76.9%; Pred. No. 0.033;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 IKGPTLRQWLKSR 14
   |:|||||:|
Db 1 IEGPTLRQWLAAK 13

```

Search completed: September 1, 2005, 16:12:12  
 Job time : 84.7482 secs

This Page Blank (uspio)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:57:33 ; Search time 13.7266 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-9

Perfect score: 97

Sequence: 1 TIKGPTLRQWLKSRHTS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 79.\*

1: p1r1.\*

2: p1r2.\*

3: p1r3.\*

4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	50.5	122	2 A10380	conserved hypothet
2	47	48.5	327	2 B71900	hypothetical prote
3	45	46.4	186	2 A90167	adenylate cyclase,
4	44	45.4	161	2 T06826	beta-fructofuranos
5	44	45.4	1019	2 T11560	pol polypeptide -
6	44	45.4	1058	2 S08436	pol polypeptide -
7	43	44.3	63	2 T10614	hypothetical prote
8	43	44.3	216	2 S35151	photosystem I chai
9	43	44.3	217	2 S46354	pol polypeptide -
10	43	44.3	233	2 A83862	initiation of chro
11	43	44.3	241	2 S07740	hypothetical prote
12	43	44.3	269	2 S73999	hypothetical prote
13	43	44.3	656	2 S30484	pol polypeptide -
14	43	44.3	656	2 S30483	pol polypeptide -
15	43	44.3	1583	2 S59644	slister chromatid c
16	43	43.8	275	2 ACO189	probable exported
17	42	43.3	114	2 E71171	hypothetical prote
18	42	43.3	171	2 T20567	hypothetical prote
19	42	43.3	416	2 D72456	probable glutamyl-
20	42	43.3	438	2 S71157	cytochrome c bioge
21	42	43.3	519	2 C86160	hypothetical prote
22	42	43.3	581	2 T12095	beta-fructofuranos
23	42	43.3	617	2 S75447	proline-CRNA ligas
24	42	43.3	1061	1 GNLJG4	HIV-1 retropepsin
25	42	43.3	2108	1 H70819	probable polyketid
26	41.5	42.8	877	2 T03098	p97 protein - Toxo
27	41	42.3	70	2 T06920	ribosomal protein
28	41	42.3	153	2 A87524	hypothetical prote
29	41	42.3	178	2 A82743	hypothetical prote

30	41	42.3	188	2 C82863	recombinase Xfa001
31	41	42.3	203	2 T40206	hypothetical prote
32	41	42.3	228	2 S76936	hypothetical prote
33	41	42.3	257	2 E70429	tRNA guanine-N1 me
34	41	42.3	264	2 AG2095	hypothetical prote
35	41	42.3	306	2 T06601	UTP-glucose-1-phos
36	41	42.3	380	2 B47029	methylase Llap1 -
37	41	42.3	397	2 G69449	tryptophan synthas
38	41	42.3	419	1 S29127	carboxypeptidase A
39	41	42.3	472	2 F70876	probable pap3 pro
40	41	42.3	622	2 S35122	site-specific DNA-
41	41	42.3	1039	2 S46347	pol polypeptide -
42	41	42.3	1123	2 T51517	telomerase reverse
43	41	42.3	1140	2 S73786	hypothetical prote
44	41	42.3	1191	2 T31081	hypothetical prote
45	41	42.3	1345	2 T13423	hypothetical prote
46	40	41.2	158	2 D72305	hypothetical prote
47	40	41.2	185	2 T49611	hypothetical prote
48	40	41.2	200	2 T23485	probable glutathio
49	40	41.2	207	2 T37464	probable glutathio
50	40	41.2	236	1 T24385	dihydropteridine r
51	40	41.2	242	1 C70895	hypothetical prote
52	40	41.2	249	2 C70895	ABC transporter, A
53	40	41.2	262	2 B83126	probable transcrip
54	40	41.2	306	2 T45453	UTP-glucose-1-phos
55	40	41.2	336	2 E72389	hypothetical prote
56	40	41.2	344	2 E84377	hypothetical prote
57	40	41.2	349	2 B97912	protein export (im
58	40	41.2	395	2 S40171	phosphoglucose 4,6-de
59	40	41.2	419	1 CPRTA	carboxypeptidase A
60	40	41.2	505	2 T19971	hypothetical prote
61	40	41.2	506	2 T19973	hypothetical prote
62	40	41.2	521	2 T01923	hypothetical prote
63	40	41.2	527	2 B64633	glucan 1,4-alpha-g
64	40	41.2	612	2 JQ1346	conserved hypothet
65	40	41.2	978	2 B89971	pol polypeptide -
66	40	41.2	1009	2 S28081	pol polypeptide -
67	40	41.2	1009	2 S44621	C50C3.2 protein -
68	40	41.2	1034	1 GNLJCA	HIV-1 retropepsin
69	40	41.2	1035	1 GNLJG6	HIV-1 retropepsin
70	40	41.2	1036	1 GNLJG2	HIV-1 retropepsin
71	40	41.2	1040	2 T08190	hypothetical prote
72	40	41.2	1054	1 GNLJG5	HIV-1 retropepsin
73	40	41.2	1055	1 GNLJST	HIV-1 retropepsin
74	40	41.2	1055	1 S53092	pol polypeptide -
75	40	41.2	1056	1 GNLJG3	HIV-1 retropepsin
76	40	41.2	1712	1 CGHU28	collagen alpha 2(I
77	40	41.2	2609	2 T40399	probable transport
78	39.5	40.7	642	2 A83258	probable soluble 1
79	39.5	40.7	722	2 T37970	probable G2-specif
80	39.5	40.7	1758	2 T34393	hypothetical prote
81	39	40.2	195	2 T36141	probable nicotinam
82	39	40.2	220	2 AC0318	uracil DNA glycosy
83	39	40.2	240	2 S73922	phosphoribosylform
84	39	40.2	240	2 AE1145	hypothetical prote
85	39	40.2	314	2 H90638	nitrate reductase
86	39	40.2	318	2 S17197	hypothetical prote
87	39	40.2	320	2 H85489	probable transposa
88	39	40.2	327	2 E82277	glyceraldehyde-3-p
89	39	40.2	331	2 B48445	short-chain alchoh
90	39	40.2	336	2 A47542	dtDP-glucose 4,6-d
91	39	40.2	349	2 F86649	probable activator
92	39	40.2	373	2 D64729	hypothetical prote
93	39	40.2	410	2 H86290	hypothetical prote
94	39	40.2	533	2 C83658	adhesin AP65-1 pre
95	39	40.2	567	2 S69778	DNA43 protein - ye
96	39	40.2	571	2 S48384	hypothetical prote
97	39	40.2	615	2 T15575	E1 protein - human
98	39	40.2	644	2 W1LW58	erythrocyte membra
99	39	40.2	721	2 A39707	microtubule-associ
100	39	40.2	721	2 A33319	

## ALIGNMENTS

## RESULT 1

AI0380 conserved hypothetical protein YP03137 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AI0380  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiplall, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001  
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A/Reference number: AB0001, M01D:21470413; PMID:11586360  
A/Accession: AI0380

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-122 <KUR>  
A/Cross-references: UNIPROT:Q8ZC84; GB:AL590842; PIDN:CAC92372.1; PID:G15981075; GSPDB:C  
C/Genetics:  
A/Gene: YP03137  
C/Superfamily: Escherichia coli ybaJ protein

Query Match 50.5%; Score 49; DB 2; Length 122;  
Best Local Similarity 53.3%; Pred. No. 1.3;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRWLKSRH 16  
DB 92 INDELRWQKXEH 106

## RESULT 2

B71900 hypothetical protein jhp0682 - Helicobacter pylori (strain J99)

C/Species: Helicobacter pylori  
A/Variety: strain J99  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: B71900  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Lives, C.; Gibson, R.; Weiberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A/Reference number: A71800; M01D:99120557; PMID:9923682  
A/Accession: B71900

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-327 <ARN>  
A/Cross-references: UNIPROT:Q9ZL98; GB:AE001500; GB:AE001439; NID:G4155238; PIDN:AAD0627  
A/Experimental source: strain J99  
C/Genetics:  
A/Gene: jhp0682  
C/Superfamily: conserved hypothetical protein HI0176

Query Match 48.5%; Score 47; DB 2; Length 327;  
Best Local Similarity 50.0%; Pred. No. 8;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRWLKSRH 16  
DB 103 SVKEPTLVDMKSNY 118

## RESULT 3

A90167 adenylate cyclase, cyab-type, probable (cyab) [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 16-Aug-2004  
C/Accession: A90167  
R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.U.; Chan, J.; Jung, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.; Jurek, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.  
A/Reference number: A99139  
A/Accession: A90167  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-186 <KUR>  
A/Cross-references: UNIPROT:Q980N7; GB:AE006641; NID:G13813390; PIDN:AAK40592.1; GSPDB:C  
C/Genetics:  
A/Gene: CYAB  
C/Superfamily: Thermophilic adenylate cyclase, Cyab type

Query Match 45.4%; Score 45; DB 2; Length 186;  
Best Local Similarity 55.6%; Pred. No. 9;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TIKGPTLRWLKSRH 18  
DB 65 TYKGPTLHSSLKAREIS 82

## RESULT 4

T06826 beta-fructofuranosidase (FC 3.2.1.26) II - garden pea (fragment)

N/Alternate names: cell wall invertase II  
C/Species: Pisum sativum (garden pea)  
C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T06826  
R:Buchner, P.  
submitted to the EMBL Data Library, December 1996  
A/Reference number: Z15838

A/Accession: T06826  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-161 <BUC>  
A/Cross-references: UNIPROT:P93490; EMBL:Z83339; PIDN:CAB05954.1  
C/Superfamily: beta-fructofuranosidase  
C/Keywords: glycosidase; hydrolase

Query Match 45.4%; Score 44; DB 2; Length 161;  
Best Local Similarity 46.7%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRWLKSRH 16  
DB 76 VSDPFLREWKSPEN 90

## RESULT 5

T11560 pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)

C/Species: simian immunodeficiency virus SIVsm  
A/Variety: strain E543  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T11560  
R:Hirsch, V.M.; Adger-Johnson, D.; Camball, B.; Goldstein, S.; Brown, C.; Elkins, W.R.; J. Virol. 71, 1608-1620, 1997  
A/Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodefici  
A/Reference number: Z17285; M01D:97151152; PMID:8995688  
A/Accession: T11560

A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1019 <HIR>  
A/Cross-references: UNIPROT:P89154; EMBL:U72748; NID:G1695908; PIDN:AA056559.1; PID:G16  
C/Genetics:  
A/Gene: pol  
C/Superfamily: pol polyprotein  
C/Keywords: AIDS; immunodeficiency

Query Match 45.4%; Score 44; DB 2; Length 1019;  
Best Local Similarity 61.5%; Pred. No. 82;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTRLQWLKSRRE 15  
 :|||:|:|:  
 Db 184 EGPRLQWPLSKRE 196

## RESULT 6

pol polypeptide - human immunodeficiency virus type 2 D205 (fragment)  
 C:Species: human immunodeficiency virus type 2 D205, HIV-2 D205  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: S08436  
 R:Detrich, U.; Adamki, M.; Kreutz, R.; Seipp, A.; Kuehn, H.; Ruebsamen-Waigmann, H.  
 Nature 342, 948-950, 1989  
 A:Title: A highly divergent HIV-2-related isolate.  
 A:Reference number: S08434; MUID:90081881; PMID:2594088  
 A:Accession: S08436  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1058 <DIE>  
 A:Cross-references: UNIPROT:P15833; EMBL:X16109  
 A:Note: this sequence was submitted to the EMBL Data Library, Aug-1989  
 C:Genetics:  
 A:Gene: pol  
 C:Superfamily: pol polypeptide  
 C:Keywords: polypeptide

Query Match 45.4%; Score 44; DB 2; Length 1058;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRRE 15  
 ||:|:|:|:  
 Db 222 GPKIRQWPLSKRE 233

## RESULT 7

hypothetical protein 12L - Molluscum contagiosum virus 1  
 N:Alternate names: MC0012L  
 C:Species: Molluscum contagiosum virus 1  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
 C:Accession: T30614  
 R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
 Science 275, 813-816, 1996  
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
 A:Reference number: 220876; MUID:96325459; PMID:8670425  
 A:Accession: T30614  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-63 <SEN>  
 A:Cross-references: UNIPROT:Q98183; EMBL:U60315; PIDN:AAC55140.1  
 C:Genetics:  
 A:Note: MC012L

Query Match 44.3%; Score 43; DB 2; Length 63;  
 Best Local Similarity 41.2%; Pred. No. 5.8;  
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTRLQWLKSRREHTS 18  
 :|||:|:|:|:  
 Db 41 VLGETLRWRSKRNTA 57

## RESULT 8

S35151  
 photosystem I chain XI precursor - spinach  
 C:Species: Spinacia oleracea (spinach)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S35151; S14446  
 R:Pileger, K.; Oelmueller, R.; Hermann, R.G.  
 Plant Mol. Biol. 22, 703-709, 1993  
 A:Title: Isolation and characterization of cDNA clones encoding a 18.8 kDa polypeptide,  
 A:Reference number: S35151; MUID:93344519; PMID:8343606

A:Accession: S35151  
 A:Molecule type: mRNA  
 A:Residues: 1-216 <FRI>  
 A:Cross-references: UNIPROT:Q41385; EMBL:X64445; NID:G396274; PIDN:CAA45775.1; PID:G396  
 A:Experimental source: seed  
 R:Ikeuchi, M.; Inoue, Y.  
 FEBS Lett. 280, 332-334, 1991

A:Title: Two new components of 9 and 14 kDa from spinach photosystem I complex.  
 A:Reference number: S14316; MUID:91192162; PMID:2013332  
 A:Accession: S14446  
 A:Molecule type: protein  
 A:Residues: 158-175, 'X', 177-178 <IKS>  
 C:Genetics:  
 A:Gene: psal  
 A:Function:  
 C:Description: this protein is a component of photosystem I which catalyzes the light-  
 C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid; transmembrane protein  
 F:1-47/Domain: transit peptide (chloroplast) #status predicted <TMP>  
 F:48-216/Product: photosystem I chain XI #status predicted <MAT>  
 F:131-153/Domain: transmembrane #status predicted <TM1>  
 F:187-209/Domain: transmembrane #status predicted <TM2>

Query Match 44.3%; Score 43; DB 2; Length 216;  
 Best Local Similarity 52.9%; Pred. No. 22;  
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 IKGPTRLQWLKSRREHTS 18  
 |||:|:|:|:  
 Db 27 ISGPALRGFPSPRRHTS 43

## RESULT 9

S46354  
 pol polypeptide - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)  
 C:Species: simian immunodeficiency virus SIVagm  
 A:Variety: isolate SABD37  
 C:Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 26-Aug-1999  
 C:Accession: S46354  
 R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.,  
 EMBO J. 13, 2935-2947, 1994  
 A:Title: Mosaic genome structure of simian immunodeficiency virus from West African grc  
 A:Reference number: S46335; MUID:94298785; PMID:8026477  
 A:Accession: S46354  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-217 <JIN>  
 A:Cross-references: EMBL:U04018; NID:G466250; PIDN:AAA21512.1; PID:G466251  
 A:Experimental source: isolate SABD37; baboon monkey  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993  
 C:Genetics:  
 A:Gene: pol  
 C:Superfamily: pol polypeptide  
 C:Keywords: polypeptide

Query Match 44.3%; Score 43; DB 2; Length 217;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRRE 15  
 |||:|:|:|:  
 Db 87 GPKLRQWPLSKRE 98

## RESULT 10

A83862  
 initiation of chromosome replication dnaB [imported] - Bacillus halodurans (strain C-12)  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: A83862  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132

A/Accession: A83862  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-233 <STO>  
A/Cross-references: UNIPROT:Q9KC77; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA054  
A/Experimental source: strain C-125  
C/Genetics:  
A/Gene: dhad

Query Match 44.3%; Score 43; DB 2; Length 233;  
Best Local Similarity 43.8%; Pred. No. 24;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRREHT 17  
Db 144 IEGETLSMWIDQOQHT 159

## RESULT 11

S07740

hypothetical protein 8 - Paramaecium tetraurelia mitochondrion

C/Species: mitochondrion Paramaecium tetraurelia

C/Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004

C/Accession: S07740

R/Pitchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummings,

Nucleic Acids Res. 18, 173-180, 1990

A/Title: Nucleotide sequence of the mitochondrial genome of Paramaecium.

A/Reference number: S07725; MUID:90174913; PMID:2308823

A/Accession: S07740

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-241 &lt;PRI&gt;

A/Cross-references: UNIPROT:P15609; EMBL:X15917; NID:g13356; PID:g578757

C/Genetics:

A/Genome: mitochondrion

A/Genetic code: SGC6

A/Start codon: ATT

C/Superfamily: mitochondrial ribosomal protein S18, paramaecium type

C/Keywords: mitochondrion

Query Match 44.3%; Score 43; DB 2; Length 241;  
Best Local Similarity 41.2%; Pred. No. 25;  
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRREHTS 18  
Db 26 VKGPTIEKFLKRFYNA 42

## RESULT 12

S73999

hypothetical protein yaac homolog VXDSP7\_orf269 - Mycoplasma pneumoniae (strain ATCC 29

N/Alternate names: hypothetical protein VXDSP7\_orf269

C/Species: Mycoplasma pneumoniae

A/Variety: ATCC 29342

C/Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

C/Accession: S73999

R/Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkil, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A/Reference number: S73327; MUID:97105885; PMID:8948633

A/Accession: S73999

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-269 &lt;HMT&gt;

A/Cross-references: UNIPROT:P75587; EMBL:AE000062; GB:U00089; NID:g1674373; PIDN:AA59632

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C/Genetics:

A/Genetic code: SGC3

C/Superfamily: uncharacterized conserved protein HI0963

Query Match 44.3%; Score 43; DB 2; Length 269;  
Best Local Similarity 58.3%; Pred. No. 28;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLK 12  
Db 124 TLSSSTRQWLK 135

## RESULT 13

S30484

pol polyprotein - human immunodeficiency virus type 2

C/Species: human immunodeficiency virus type 2, HIV-2

C/Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001

C/Accession: S30484

R/Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barcane, J.; Hanson, A.P.; Greene, B.M.;

submitted to the EMBL Data Library, December 1992

A/Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa

A/Reference number: S30460

A/Accession: S30484

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-656 &lt;GAO&gt;

A/Cross-references: EMBL:M87114

C/Superfamily: pol polyprotein

Query Match 44.3%; Score 43; DB 2; Length 656;  
Best Local Similarity 66.7%; Pred. No. 74;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRRE 15  
Db 31 GPTLRQWLKSRRE 42

## RESULT 14

S30483

pol polyprotein - human immunodeficiency virus type 2

C/Species: human immunodeficiency virus type 2, HIV-2

C/Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001

C/Accession: S30483

R/Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barcane, J.; Hanson, A.P.; Greene, B.M.;

submitted to the EMBL Data Library, December 1992

A/Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa

A/Reference number: S30460

A/Accession: S30483

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-656 &lt;GAO&gt;

A/Cross-references: EMBL:M87111

C/Superfamily: pol polyprotein

Query Match 44.3%; Score 43; DB 2; Length 656;  
Best Local Similarity 66.7%; Pred. No. 74;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRRE 15  
Db 31 GPTLRQWLKSRRE 42

## RESULT 15

S59644

sister chromatid cohesion molecule Mts4p - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 14-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004

C/Accession: T38603; T43392; S59644

R/Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, July 1995

A/Reference number: Z21731

A/Accession: T38603

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1583 &lt;DE2&gt;

A/Cross-references: UNIPROT:Q09725; EMBL:Z50113; NID:g914878; PIDN:CAA90463.1; PID:g914

A;Experimental source: strain 972b-; cosmid c31A2  
R;Ruruya, K.; Takahashi, K.; Yanagida, M.  
submitted to the EMBL Data Library, August 1998  
A;Description: Faithful anaphase is ensured by Msa4, a sister chromatid cohesion molecule  
A;Reference number: 222478  
A;Accession: T43392  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1583 <FUR>  
A;Cross-references: EMBL:AB016866; PIDN:CAB19489.1  
C;Genetics:  
A;Gene: msa4; SPAC31A2.05c  
A;Map position: 1  
A;introns: 33/1; 98/2; 543/3; 699/3; 1294/2; 1339/3; 1558/3  
Query Match 44.3%; Score 43; DB 2; Length 1583;  
Best Local Similarity 46.2%; Pred. No. 1.9e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
OY 4 GPTLRQWLKSRH 16  
DB 1483 GPTTGMKKLDH 1495  
RESULT 16  
AC0189  
probable exported protein YPO1551 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AC0189  
R;Perkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AC0189  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-275 <KUR>  
A;Cross-references: UNIPROT:Q8ZFX2; GB:AL590842; PIDN:CAC90374.1; PID:G15979594; GSPDB:C  
C;Genetics:  
A;Gene: YPO1551  
Query Match 43.8%; Score 42.5; DB 2; Length 275;  
Best Local Similarity 47.6%; Pred. No. 34;  
Matches 10; Conservative 1; Mismatches 7; Indels 3; Gaps 1;  
OY 1 TIKGPT--LRQWLKSRH 18  
DB 133 TVAGKTMALEQWLHOLPH 153  
RESULT 17  
E71171  
hypothetical protein PH0569 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Jul-2004  
C;Accession: E71171  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekit  
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: E71171  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-114 <KAW>  
A;Cross-references: UNIPROT:O58304; GB:AP000002; NID:G3236129; PIDN:BAA29658.1; PID:G325  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH0569

Query Match 43.3%; Score 42; DB 2; Length 114;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 2 IKGPTLRQWLK 12  
DB 47 VKGDTLKVWLK 57  
RESULT 18  
T20567  
hypothetical protein F08A10.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T20567  
R;Kershaw, J.  
submitted to the EMBL Data Library, June 1996  
A;Reference number: Z19293  
A;Accession: T20567  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-171 <WIL>  
A;Cross-references: UNIPROT:Q19186; EMBL:Z75534; PIDN:CAA9825.1; GSPDB:GN00019; CESP:  
C;Genetics:  
A;Experimental source: clone F08A10  
A;Gene: CESP.F08A10.1  
A;Map position: 1  
A;introns: 26/3; 64/1; 91/3; 118/3  
Query Match 43.3%; Score 42; DB 2; Length 171;  
Best Local Similarity 57.1%; Pred. No. 25;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
OY 4 GPTLRQWLKSRH 17  
DB 103 GPSLRPFLNSGNHT 116  
RESULT 19  
D72456  
probable glutamyl-tRNA reductase APE2296 - Aeropyrum pernix  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: D72456  
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Jin-no, K.; Take  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
DNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: D72456  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-416 <KAW>  
A;Cross-references: UNIPROT:Q9Y9J2; DBJ:AP000064; NID:G5105945; PIDN:BAA81308.1; PID:  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE2296  
C;Superfamily: glutamyl-tRNA reductase  
Query Match 43.3%; Score 42; DB 2; Length 416;  
Best Local Similarity 57.1%; Pred. No. 65;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
OY 2 IKGPTLRQWLKSR 15  
DB 105 VLGQVRRAWLKSR 118  
RESULT 20  
S71157  
cytochrome c biogenesis protein 454 - evening primrose mitochondrion  
C;Species: mitochondrion Oenothera lamarckiana (evening primrose)

C>Date: 28-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
C/Accession: S71157, S55283, S42984  
R/Gruska, I.; Jekabsons, W.; Schuster, W.  
Mol. Gen. Genet. 247, 529-536, 1995  
A>Title: Genochera mitochondrial orf454, a gene involved in cytochrome c biogenesis cor  
A/Reference number: S55283; MUID:95327048; PMID:7603431  
A/Accession: S71157  
A/Molecule type: mRNA  
A/Residues: 1-438 <GRU>  
A/Cross-references: UNIPROT:Q35213; EMBL:X78036  
A/Note: differences are due to RNA editing; premature stop codon is due to RNA editing  
A/Accession: S55283  
A/Molecule type: DNA  
A/Residues: 1-16, 'PR', 19-34, 'P', 36-39, 'SS', 42-48, 'P', 50, 'PS', 53, 'P', 55-108, 'U', 110-130, '  
A/Cross-references: EMBL:X78036; NID:9459536; PIDN:CAAS4966.1; PID:9459537  
A/Genetics:  
A/Genome: mitochondrion  
A/Introns: 256/2  
C/Keywords: mitochondrion; RNA editing

Query Match 43.3%; Score 42; DB 2; Length 438;  
Best Local Similarity 50.0%; Pred. No. 68;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 7 LROMLKSRHETS 18  
|:|:|:|:|:  
Db 342 LHRVWKNRHHNN 353

RESULT 21  
C86160  
hypothetical protein F22D16.3 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: C86160  
R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Matti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: C86160  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-519 <STO>  
A/Cross-references: UNIPROT:Q9SRV9; GB:AE005172; NID:56056405; PIDN:AAE02869.1; GSPDB:GN  
C/Genetics:  
A/Map position: 1

Query Match 43.3%; Score 42; DB 2; Length 519;  
Best Local Similarity 40.0%; Pred. No. 82;  
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 2 IKGPTLROMLKSRH 16  
|:|:|:|:|:|:  
Db 498 IGRHTWEMWAKON 512

RESULT 22  
T12095  
beta-fructofuranosidase (EC 3.2.1.26), cell wall - fava bean  
N/Alternate names: cell wall invertase II  
C/Species: Vicia faba (fava bean)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T12095  
R/Weder, H.; Borisjuk, L.; Heim, U.; Buchner, P.; Mobus, U.  
Plant Cell 7, 1833-1846, 1995  
A>Title: Seed coat-associated invertases of fava bean control both unloading and storage

```

A:Reference number: Z17416; MUID:96093423; PMID:8535137
A:Accession: J12095
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-581 <MBE>
A:Cross-references: UNIPROT:Q43856; EMBL:Z35163; NID:g861156; PIDN:CAA84527.1; PID:g861
A:Experimental source: strain minor; cultivar F1bdo; seed coat; clone VECWINV2
C:Genetics:
A:Gene: CWINV2
C:Function:
A:Description: hydrolyzes terminal non-reducing beta-D-fructofuranoside residues in bet
A:Superfamily: beta-fructofuranosidase
C:Keywords: cell wall; glycoprotein; glycosidase; hydrolase

Query Match 43.3% Score 42; DB 2; Length 581;
Best Local Similarity 58.3%; Pred. No. 93;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PTLRQWLKSRHH 16
Db 160 PFLREWIKSPEN 171

RESULT 23
S75447
Proline-tRNA ligase - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1425
C:Species: Synechocystis sp.
A:Variatey: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S75447
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75447
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-617 <KAN>
A:Cross-references: UNIPROT:P73942; EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAJ1801
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: proS
A:Start codon: GTG
C:Superfamily: proline-tRNA ligase

Query Match 43.3% Score 42; DB 2; Length 617;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLRQWLKSRHH 18
Db 604 PTLTAMTAEKKTTS 617

RESULT 24
GNLJG4
HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (African green monkey
N:contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA pol
C:Species: simian immunodeficiency virus, SIV
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
C:Accession: B30045
R:Fukusawa, M.; Miura, T.; Hasegawa, A.; Morikawa, S.; Tsujimoto, H.; Miki, K.; Kitamura
Nature 333, 457-461, 1988
A:Title: Sequence of simian immunodeficiency virus from African green monkey, a new mem
A:Reference number: A30045; MUID:88232506; PMID:3374586
A:Accession: B30045
A:Molecule type: DNA
A:Residues: 1-1061 <PUK>
A:Cross-references: EMBL:X07805; NID:g61748; PID:g1335593
C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, r

```

C:Genetics:  
 A:Gene: pol  
 C:Superfamily: pol polypeptide  
 C:Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polypeptide; reverse  
 F:11-210/Product: retropepsin #status predicted <RFP>  
 F:114/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 43.3%; Score 42; DB 1; Length 1061;  
 Best Local Similarity 53.8%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 3 KGPTRLQWLKSRRE 15  
 Db 229 RGPVTRQWLKSRRE 241

## RESULT 25

H70819  
 Probable polyketide synthase - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: H70819

R:Cole, S.T.; Broch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulstrom, J.B.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: H70819  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2108 <COL>

A:Cross-references: UNIPROT:O53901; GB:AL022000; GB:AL123456; NID:g3261541; PIDN:CAA1759  
 A:Experimental source: strain H37RV  
 C:Genetics8:

A:Gene: pks5  
 C:Superfamily: mycosceroid acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I hcnase homology; [acyl-carrier-protein] S-malonyltransferase homology  
 C:Keywords: carrier protein

F:36-434/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>  
 F:546-826/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>  
 F:1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>  
 F:1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>  
 F:2029-2094/Domain: acyl carrier protein homology <ACP1>

Query Match 43.3%; Score 42; DB 2; Length 2108;  
 Best Local Similarity 46.7%; Pred. No. 3.8e+02;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 2 IKGPTRLQWLKSRRE 16  
 Db 1047 VDGAEVRQWLKSRRE 1061

## RESULT 26

T03098  
 p97 protein - Toxoplasma gondii  
 C:Species: Toxoplasma gondii  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
 C:Accession: T03098

R:Matsumura, T.; Kasper, L.  
 Mol. Biochem. Parasitol. 90, 403-413, 1997  
 A:Title: Molecular analysis and characterization of a protein involved in the replicatio  
 A:Reference number: Z14838; MUID:98135655; PMID:9476788

A:Accession: T03098  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-877 <MAT>

A:Cross-references: UNIPROT:O15644; EMBL:AF005059; NID:g2581824; PIDN:AA47857.1; PID:g2  
 C:Function:  
 A:Description: involved in replication of intracellular Toxoplasma gondii

C:Superfamily: Toxoplasma gondii p97 protein

Query Match 42.8%; Score 41.5; DB 2; Length 877;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Oy 4 GP---TLQWLKSRRE 16  
 Db 72 GVPNTRQWLQQRNEH 87

## RESULT 27

T06920  
 ribosomal protein l28 - Cyanophora paradoxa cyanelle  
 C:Species: cyanelle Cyanophora paradoxa  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T06920

R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.  
 submitted to the EMBL Data Library, July 1995

A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.  
 A:Reference number: Z15840

A:Accession: T06920  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Residues: 1-70 <STI>  
 A:Cross-references: UNIPROT:P48129; EMBL:U30821; NID:g1016083; PIDN:AAA81263.1; PID:g10

A:Experimental source: strain Pflingsheim LB555  
 C:Genetics:  
 A:Gene: rpl28  
 A:Superfamily: Escherichia coli ribosomal protein l28  
 C:Keywords: cyanelle; ribosome

Query Match 42.3%; Score 41; DB 2; Length 70;  
 Best Local Similarity 50.0%; Pred. No. 13;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 2 IKGPTRLQWLKSRRE 17  
 Db 38 IWSPTLQWLKSRRE 53

## RESULT 28

A97524  
 hypochemical protein AGR\_C\_2500 [imported] - Agrobacterium tumefaciens (strain C58, Cer  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: A97524

R:Goodner, B.; Hinkle, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Gurollo, B.; Goldman  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B  
 Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium t  
 A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: A97524  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-153 <KUR>  
 A:Cross-references: UNIPROT:Q8UPF3; GB:AB007869; PIDN:AAK87146.1; PID:g15156416; GSPDB  
 C:Genetics:  
 A:Gene: AGR\_C\_2500

A:Map position: circular chromosome  
 C:Superfamily: 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase; 2-a

Query Match 42.3%; Score 41; DB 2; Length 153;  
 Best Local Similarity 37.5%; Pred. No. 32;  
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 2 IKGPTRLQWLKSRRE 17  
 Db 119 VKGRPVQWLQQRDRS 134

## RESULT 29



AB2743  
 hypothetical protein foik [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: AB2743  
 R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McNeill  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 eter, E.W.  
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:2160850; PMID:11743193  
 A:Accession: AB2743  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-178 <KUR>  
 A:Cross-references: UNIPROT:Q8FPP3; GB:AE008688; PIDD:RAL42360.1; PID:G17739767; GSPDB:C  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: foik  
 A:Map position: circular chromosome  
 C:Superfamily: 2-amino-4-hydroxy-6-hydroxymethylidihydroperidine pyrophosphokinase; 2-am  
 Query Match 42.3%; Score 41; DB 2; Length 178;  
 Best Local Similarity 37.5%; Pred. No. 37;  
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 Oy 2 IKGPTLROWLKSREHT 17  
 Db 144 VKGKPVQWLOQADRS 159

RESULT 30  
 C82863  
 recombinase XFA019 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C:Accession: C82863  
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; PMID:20365717; PMID:10910347  
 A>Note: for a complete list of authors see reference number A59328 below  
 A:Accession: C82863  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-188 <SIM>  
 A:Cross-references: UNIPROT:Q9PH16; GB:AE003851; NID:g9112238; PIDD:JAR85588.1; GSPDB:GN  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Aencio, M.; Alvarenga, R.; A  
 Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
 submitted to Genbank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 ; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tsubako, M.H.; Vailada, H.; Van Sluys, W.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XFA019  
 A:Genome: plasmid  
 A>Note: Plasmid pXFS.1  
 C:Superfamily: transposase repressor  
 Query Match 42.3%; Score 41; DB 2; Length 188;  
 Best Local Similarity 53.8%; Pred. No. 39;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 5 PTLROWLKSREHT 17  
 Db 176 PTLRWLPASAH 188

RESULT 31  
 T40206  
 hypothetical protein SPBC31F10.03 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T40206  
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.  
 submitted to the EMBL Data Library, August 1997  
 A:Accession: T40206  
 A:Reference number: Z21913  
 A:Molecule type: DNA  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Residues: 1-203 <WOO>  
 A:Cross-references: UNIPROT:P87305; EMBL:Z97204; PIDD:CAB10080.1; GSPDB:GN00067; SPDB:S  
 A:Experimental source: strain 972h-; cosmid C31F10  
 C:Genetics:  
 A:Gene: SPDB:SPBC31F10.03  
 A:Map position: 2  
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC31F10.03  
 Query Match 42.3%; Score 41; DB 2; Length 203;  
 Best Local Similarity 46.7%; Pred. No. 43;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Oy 2 IKGPTLROWLKSREH 16  
 Db 34 IKGYRRFWRSSEH 48

RESULT 32  
 S76936  
 hypothetical protein - Synecocystis sp. (strain PCC 6803)  
 C:Species: Synecocystis sp.  
 A:Variety: PCC 6803  
 C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S76936  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.  
 o, K.; Okumura, S.; Shiino, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud  
 DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocysti  
 S.  
 A:Reference number: S74322; PMID:97061201; PMID:8905231  
 A:Accession: S76936  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-238 <KAN>  
 A:Cross-references: UNIPROT:P74728; EMBL:D90917; GB:AB001339; NID:g1653836; PIDD:BA1188  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 Query Match 42.3%; Score 41; DB 2; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 51;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Oy 2 IKGPTLROWLKSREHT 17  
 Db 13 ILAPALRWLRSGVER 28

RESULT 33  
 E70429  
 tRNA guanine-N1 methyltransferase - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
 C:Accession: E70429  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; C  
 V.  
 Nature 392, 353-358, 1998



A>Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
 A/Reference number: A70300; MUID:98198666; PMID:9537320  
 A/Accession: E70429  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-257 <AOB>  
 A/Cross-references: UNIPROT:O67463; GB:AE000742; NID:g2983658; PIDN:AA07418.1; PID:g298  
 A/Experimental source: strain VF5  
 C/Genetics:  
 A/Gene: trmD  
 C/Superfamily: tRNA (guanine-N1) methyltransferase

Query Match 42.3%; Score 41; DB 2; Length 257;  
 Best Local Similarity 33.3%; Pred. No. 55;  
 Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRH 16  
 Db 241 LSGKSFKEMLKHKH 255

RESULT 34  
 AG2095  
 hypothetical protein all2318 [imported] - *Neotoc* sp. (strain PCC 7120).  
 C/Species: *Neotoc* sp. PCC 7120  
 A/Note: *Neotoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
 C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C/Accession: AG2095  
 R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, S.; Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120  
 A/Reference number: AB1807; MUID:21595285; PMID:11759840  
 A/Accession: AG2095  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-264 <KUR>  
 A/Cross-references: UNIPROT:O8YUW2; GB:BA000019; PIDN:BA074017.1; PID:g17131410; GSPDB:G  
 A/Experimental source: strain PCC 7120  
 C/Genetics:  
 A/Gene: all2318

Query Match 42.3%; Score 41; DB 2; Length 264;  
 Best Local Similarity 38.5%; Pred. No. 57;  
 Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PTLRQWLKSRH 17  
 Db 95 PALKQWLQSKQY 107

RESULT 35  
 D70601  
 UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - *Mycobacteri*  
 C/Species: *Mycobacterium tuberculosis*  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C/Accession: D70601  
 R/Conor, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Kaul, R.; Davies, R.; Devlin, K.; Felkell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: D70601  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-306 <COL>  
 A/Cross-references: UNIPROT:O05576; GB:I294752; GB:AL123456; NID:g3261731; PIDN:CA08153.  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Gene: galU  
 C/Superfamily: Escherichia coli UTP-glucose-1-phosphate uridylyltransferase

C/Keywords: nucleotidyltransferase

Query Match 42.3%; Score 41; DB 2; Length 306;  
 Best Local Similarity 63.6%; Pred. No. 67;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSR 14  
 Db 290 GPDLRMLVLR 300

RESULT 36  
 B47029  
 methylase llaPI - phage nck202.50 (phi 50) (fragment)  
 C/Species: phage nck202.50 (phi 50)  
 C/Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 16-Aug-2004  
 C/Accession: B47029  
 R/Hill, C.; Miller, L.A.; Klenhammer, T.R.  
 J. Bacteriol. 173, 4363-4370, 1991  
 A/Title: In vivo genetic exchange of a functional domain from a type II A methylase bet  
 A/Reference number: A47029; MUID:91294179; PMID:1906061  
 A/Contents: *Lactococcus lactis*  
 A/Accession: B47029  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-380 <HIL>  
 A/Note: sequence inconsistent with the nucleotide translation  
 A/Note: sequence extracted from NCBI backbone (NCBIN:41637, NCBI:41639)  
 C/Superfamily: Modification methylase (adenine-specific), NlaIII type

Query Match 42.3%; Score 41; DB 2; Length 380;  
 Best Local Similarity 57.1%; Pred. No. 85;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSRH 17  
 Db 81 GKTFRQWLKSRH 94

RESULT 37  
 G69449  
 tryptophan synthase (EC 4.2.1.20) beta chain - *Archaeoglobus fulgidus*  
 C/Species: *Archaeoglobus fulgidus*  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C/Accession: G69449  
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, B.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Goessens, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A/Authors: Usterbach, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
 A/Reference number: A69250; MUID:98049343; PMID:9389475  
 A/Accession: G69449  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-397 <KLE>  
 A/Cross-references: UNIPROT:O28672; GB:AE000992; GB:AE000782; NID:g2689315; PIDN:AAB996  
 C/Genetics:  
 A/Gene: trpB-2  
 C/Function: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan  
 A/Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan  
 A/Pathway: tryptophan biosynthesis  
 A/Note: cofactor pyridoxal phosphate; last step in pathway  
 C/Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology  
 C/Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate; tryp  
 F.12.393/Domain: tryptophan synthase beta chain homology <TRP>  
 F.93/Active site: His #status predicted  
 F.94/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 42.3%; Score 41; DB 2; Length 397;  
 Best Local Similarity 63.6%; Pred. No. 89;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 7 LRQWLKSRHT 17  
 Db 181 LRDMVESFEHT 191

## RESULT 38

S29127  
 carboxypeptidase A (EC 3.4.17.1) CPA1 precursor - human  
 N/Alternate names: pancreatic carboxypeptidase A1  
 C/Species: Homo sapiens (man)  
 C/Date: 25-Feb-1994 #sequence revision 19-Jan-1996 #text change 09-Jul-2004  
 C/Accession: S29127; A34205; S08253; S02810; S71394; S02811  
 R/Catalysis: L.; Villlegas, V.; Pascual, R.; Aviles, F.X.; Wicker-Planquart, C.; Puigserver  
 Biochem. J. 287, 299-303, 1992  
 A/Title: cDNA cloning and sequence analysis of human pancreatic procarboxypeptidase A1.  
 A/Reference number: S29127; MUID:93038569; PMID:1417781  
 A/Accession: S29127  
 A/Molecule type: mRNA  
 A/Residues: 1-419 <CMT>  
 A/Cross-references: UNIPROT:P15085; EMBL:X67318; NID:G35329; PIDN:CAA47732.1; PID:G35330  
 R/Stewart, E.A.; Craik, C.S.; Hake, L.; Bowcock, A.M.  
 Am. J. Hum. Genet. 46, 795-800, 1990  
 A/Title: Human carboxypeptidase A identifies a BgIII RFLP and maps to 7q31-qter.  
 A/Reference number: A34205; MUID:90196012; PMID:1969228  
 A/Accession: A34205  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 330-396 <STE>  
 A/Note: the authors translated the codon CTG for residue 391 as Val  
 R/Mouillard, M.; Michon, T.; Kerfelec, B.; Chapuis, C.  
 FEBS Lett. 261, 179-183, 1990  
 A/Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase A  
 A/Reference number: S08253; MUID:90169111; PMID:2307232  
 A/Accession: S08253

A/Molecule type: protein  
 A/Residues: 17-43; 'XXK', 114-135 <MOU>  
 R/Pascual, R.; Burgos, F.J.; Salva, M.; Soriano, F.; Mendez, E.; Aviles, F.X.  
 Eur. J. Biochem. 179, 609-616, 1989  
 A/Title: Purification and properties of five different forms of human procarboxypeptidase A  
 A/Reference number: S02809; MUID:91505096; PMID:2920728  
 A/Accession: S02810  
 A/Molecule type: protein  
 A/Residues: 17-42 <PAS>  
 R/Laethem, R.M.; Blumenkopf, T.A.; Cory, M.; Elwell, L.; Moxham, C.P.; Ray, P.H.; Walton  
 Arch. Biochem. Biophys. 332, 8-18, 1996  
 A/Title: Expression and characterization of human pancreatic preprocarboxypeptidase A1  
 A/Reference number: S71394; MUID:96400327; PMID:8806703  
 A/Accession: S71394  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-419 <LAE>  
 C/Genetics:  
 A/Gene: GDB:CPA1; CPA  
 A/Cross-references: GDB:120597; OMIM:114850  
 A/Map position: 7q32-7qter  
 C/Superfamily: carboxypeptidase  
 C/Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein digestion; zinc  
 F.1-16/Domain: signal sequence #status predicted <SIG>  
 F.11-110/Domain: activation peptide #status predicted <ACP>  
 F.111-419/Product: carboxypeptidase A isozyme 1 #status predicted <MAT>  
 F.119,182,306/Binding site: zinc (His, Glu, His) #status predicted  
 F.248-271/Disulfide bonds: #status predicted  
 F.358,380/Active site: Tyr, Glu #status predicted

Query Match 42.3%; Score 41; DB 1; Length 419;  
 Best Local Similarity 53.8%; Pred. No. 94;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 LRQWLKSRHTS 18  
 Db 232 LRDMVESFEHT 244

## RESULT 39

F70876  
 probable papA3 protein - Mycobacterium tuberculosis (strain H37RV)  
 C/Species: Mycobacterium tuberculosis  
 C/Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text change 09-Jul-2004  
 C/Accession: F70876  
 R/Cole, S.T.; Broesch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentile, S.; Hamlin, N.; Holroyd, S.  
 Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: F70876  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-472 <COL>  
 A/Cross-references: UNIPROT:O50438; GB:AL010186; GB:AL123456; NID:G3261493; PIDN:CAA158;  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Gene: papA3

Query Match 42.3%; Score 41; DB 2; Length 472;  
 Best Local Similarity 38.9%; Pred. No. 1,1e+02;  
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18  
 Db 224 TVESPQVRAWTKFAEGTN 241

## RESULT 40

S35122  
 site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) LlaI - Lactococcus  
 N/Alternate names: type II modification methylase LlaI  
 C/Species: Lactococcus lactis  
 C/Date: 16-Apr-1997 #sequence revision 09-May-1997 #text change 09-Jul-2004  
 C/Accession: S35122; S77702; A47029  
 R/Hill, C.; Miller, L.A.; Kjaehammer, T.R.  
 J. Bacteriol. 173, 4363-4370, 1991  
 A/Title: In vivo genetic exchange of a functional domain from a type II A methylase bet  
 A/Reference number: A47029; MUID:91294179; PMID:1906061  
 A/Accession: S35122

A/Molecule type: DNA  
 A/Residues: 1-622 <HIL>  
 A/Cross-references: UNIPROT:P3516; EMBL:M77136  
 A/Experimental source: bacteriophage resistance plasmid pTR2030  
 A/Note: the sequence of residues 469 and 470 is interchanged in the authors' translation.  
 A/Note: sequence extracted from NCBI backbone (NCBIN:41635, NCBIPI:41636)  
 R/Kjaehammer, T.R.  
 submitted to the EMBL Data Library, November 1994  
 A/Reference number: S77702  
 A/Accession: S77702  
 A/Molecule type: DNA  
 A/Residues: 1-248; 'G', 250-622 <KLA>  
 A/Cross-references: EMBL:U17253; NID:G639886; PIDN:AAA65073.1; PID:G639892  
 C/Genetics:  
 A/Gene: LlaI  
 A/Genome: plasmid

C/Keywords: methyltransferase; restriction modification system; S-adenosylmethionine  
 Query Match 42.3%; Score 41; DB 2; Length 622;  
 Best Local Similarity 57.1%; Pred. No. 1,4e+02;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 4 GPTLRQWLKSRHT 17  
 Db 81 GKTPQWLNERHT 94

## RESULT 41

S46347

pol polyprotein - simian immunodeficiency virus SIVagm (isolate SAB-1)  
C:Species: simian immunodeficiency virus SIVagm  
A:Variety: isolate SAB-1  
C>Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 26-Aug-1999  
C:Accession: S46347  
R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;  
EMBO J. 13, 2935-2947, 1994  
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African green  
A:Reference number: S46345; MUID:94298785; PMID:8026477  
A:Accession: S46347  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1039 <JIN>  
A:Cross-references: EMBL:U04005; NID:9466229; PIDN:AAA21505.1; PID:9466231  
A:Experimental source: isolate SAB-1; baboon monkey  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993  
C:Genetics:  
A:Gene: pol  
C:Superfamily: pol polyprotein

Query Match 42.3%; Score 41; DB 2; Length 1039;  
Best Local Similarity 58.3%; Pred. No. 2.5e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

4 GPTLRQWLKSRRE 15  
|||:|||||:  
Db 205 GPRIRQWPLSKRE 216

RESULT 42  
T51517  
telomerase reverse transcriptase - Arabidopsis thaliana  
N:Alternate names: protein F5E19\_190  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: T51517  
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
submitted to the Protein Sequence Database, August 2000  
A:Reference number: Z25394  
A:Accession: T51517  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1123 <SAT>  
A:Cross-references: UNIPROT:Q9SPU7; EMBL:AL391147  
A:Experimental source: cultivar Columbia; BAC clone F5E19  
C:Genetics:  
A:Map position: 5  
A:Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2  
A:Note: F5E19\_190

Query Match 42.3%; Score 41; DB 2; Length 1123;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

2 IKGPTLRQWLKS 13  
:|||:|||||:  
Db 200 VQPTKRWLSS 211

RESULT 43  
S73786  
hypothetical protein A19\_orf1140 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C>Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S73786  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:9710585; PMID:8946633  
A:Accession: S73786  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-1140 <HIM>  
A:Cross-references: UNIPROT:P75405; EMBL:AE000045; GB:U00089; NID:g1674145; PIDN:AA8961  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3

Query Match 42.3%; Score 41; DB 2; Length 1140;  
Best Local Similarity 53.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

4 GPTLRQWLKSRRE 16  
|||:|||||:  
Db 1116 GVTLHRWRKSRKH 1128

RESULT 44  
T31091  
hypothetical protein wdbk [imported] - Serratia marcescens  
C:Species: Serratia marcescens  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T31091  
R:Saigi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A.;  
J. Bacteriol. 181, 1883-1891, 1999  
A:Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.  
A:Reference number: Z20974; MUID:99173913; PMID:10074083  
A:Accession: T31091  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1191 <SAI>  
A:Cross-references: UNIPROT:O52484; EMBL:AF038816; NID:g2828669; PID:g2828673; PIDN:AAC  
C:Genetics:  
A:Gene: wdbk

Query Match 42.3%; Score 41; DB 2; Length 1191;  
Best Local Similarity 50.0%; Pred. No. 2.9e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

5 PTLRWLKSREHTS 18  
|||:|||||:  
Db 491 PELTQWLREBETA 504

RESULT 45  
T13423  
hypothetical protein 3088.4 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13423  
R:Murphy, D.; Harris, D.; Barrell, B.  
submitted to the EMBL Data Library, April 1999  
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A:Reference number: Z17668  
A:Accession: T13423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3345 <MUR>  
A:Cross-references: UNIPROT:O46074; EMBL:AL009195; NID:e1355203; PID:e1248585; PIDN:CAA  
C:Genetics:  
A:Map position: X  
A:Introns: 51/3; 159/1; 476/1; 526/1; 1465/1; 1826/3; 1947/3; 2081/1; 2196/3; 3007/3  
A:Note: EG:3088.4

Query Match 42.3%; Score 41; DB 2; Length 3345;  
Best Local Similarity 58.3%; Pred. No. 9e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

2 IKGPTLRQWLKS 13  
:|||:|||||:  
Db 2203 VKNPKLRQWLAS 2214

Fri Sep 2 09:00:11 2005

us-10-083-768-9.rpt

Page 12

Search completed: September 1, 2005, 16:22:54  
Job time : 16.7266 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 66.9496 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-9

Perfect score: 97

Sequence: 1 TIKGPTLRQMKSRHETS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	50.5	122	2	Q66DR4
2	49	50.5	122	2	Q82C84
3	48	49.5	126	2	P95613
4	47	48.5	327	1	Y745 HELPJ
5	47	48.5	327	1	Q750V6
6	46.5	47.9	332	1	GPD2 MYCPA
7	46	47.4	126	2	Q6YUC3
8	46	47.4	365	2	Q6C114
9	46	47.4	402	2	Q82R87
10	46	47.4	941	2	Q80UJ6
11	45	46.4	186	2	Q980N7
12	45	46.4	243	2	Q8XR40
13	45	46.4	286	2	Q8T462
14	45	46.4	302	2	Q742B3
15	45	46.4	351	2	Q7V126
16	45	46.4	789	2	Q6N3B8
17	45	46.4	1715	2	Q7P578
18	44	45.4	161	2	P93490
19	44	45.4	313	2	P90433
20	44	45.4	347	2	Q8G7M6
21	44	45.4	377	2	Q82PX5
22	44	45.4	500	2	Q6AKH3
23	44	45.4	648	2	Q9P888
24	44	45.4	1017	2	Q6VGA0
25	44	45.4	1019	1	POL_STV54
26	44	45.4	1019	2	P89154
27	44	45.4	1019	2	Q7ZBR5
28	44	45.4	1019	2	Q7ZBR7
29	44	45.4	1058	1	POL_HVZD2
30	44	45.4	1928	2	Q8D674
31	43.5	44.8	561	2	Q6MDL6

32	43.5	44.8	657	2	Q8BL07	Q8BL07 pseudomonas
33	43	44.3	63	2	Q981B3	Q981B3 molluscum c
34	43	44.3	113	2	Q8VL09	Q8VL09 uncultured
35	43	44.3	124	2	Q48532	Q48532 leptochrix
36	43	44.3	215	2	Q6AM22	Q6AM22 desulfotale
37	43	44.3	216	1	PSAL_SPIOL	Q41385 spinacia ol
38	43	44.3	217	2	Q87115	Q87115 chimpanzee
39	43	44.3	233	2	Q9KC77	Q9KC77 bacillus ha
40	43	44.3	237	1	PIRE_GLOVI	Q7HK22 gloeobacter
41	43	44.3	238	2	Q8XSR8	Q8XSR8 ralsstonia s
42	43	44.3	241	1	YMO8_PARTE	P15609 parametium
43	43	44.3	269	1	RIBF_MYCPN	P75587 mycoplasma
44	43	44.3	340	2	Q8UN03	Q8UN03 chimpanzee
45	43	44.3	340	2	Q8UN04	Q8UN04 chimpanzee
46	43	44.3	349	2	Q7SKX8	Q7SKX8 human immun
47	43	44.3	349	2	Q7SKX9	Q7SKX9 human immun
48	43	44.3	352	1	ID12_PYRAE	Q8ZYF6 pyrobaculum
49	43	44.3	369	2	Q830B7	Q830B7 enterococcu
50	43	44.3	396	2	Q90PU1	Q90PU1 chimpanzee
51	43	44.3	450	2	Q9XHE9	Q9XHE9 prunus arme
52	43	44.3	454	2	Q897U6	Q897U6 clostridium
53	43	44.3	472	2	Q9WH29	Q9WH29 human immun
54	43	44.3	527	1	TP6B_PYRAE	Q8VWZ6 pyrobaculum
55	43	44.3	536	2	Q8VYF2	Q8VYF2 arabidopsis
56	43	44.3	537	2	Q946D4	Q946D4 arabidopsis
57	43	44.3	609	2	Q856X8	Q856X8 mycobacteri
58	43	44.3	900	2	Q8GHS7	Q8GHS7 pseudomonas
59	43	44.3	986	2	Q57059	Q57059 chimpanzee
60	43	44.3	1005	2	Q6Y8X5	Q6Y8X5 human immun
61	43	44.3	1022	1	POL_STVSP	P19505 simian immun
62	43	44.3	1022	2	Q90317	Q90317 chimpanzee
63	43	44.3	1022	2	Q87956	Q87956 chimpanzee
64	43	44.3	1022	2	Q87965	Q87965 chimpanzee
65	43	44.3	1022	2	Q88135	Q88135 chimpanzee
66	43	44.3	1022	2	Q89620	Q89620 chimpanzee
67	43	44.3	1056	2	Q04097	Q04097 chimpanzee
68	43	44.3	1057	1	POL_STVAI	Q02836 simian immun
69	43	44.3	1150	2	P90246	P90246 feline immu
70	43	44.3	1150	2	Q64F60	Q64F60 feline immu
71	43	44.3	1226	2	Q6H0K6	Q6H0K6 human immun
72	43	44.3	1583	1	MIS4_SCHHO	Q09723 schizosacch
73	43	44.3	1896	1	VITI_PERAM	Q9UBM0 periplaneta
74	42.5	43.8	88	2	Q6T515	Q6T515 oryza sativ
75	42.5	43.8	275	2	Q66C47	Q66C47 yersinia ps
76	42.5	43.8	275	2	Q82EX2	Q82EX2 yersinia pe
77	42.5	43.8	307	2	Q761R3	Q761R3 sulfolobus
78	42.5	43.8	595	2	Q6N169	Q6N169 corynebacte
79	42.5	43.8	2281	2	Q6RKL2	Q6RKL2 gibberella
80	42	43.3	114	2	Q58304	Q58304 pyrococcus
81	42	43.3	173	2	Q8D9X7	Q8D9X7 vibrio vuln
82	42	43.3	189	2	Q9MFE1	Q9MFE1 beta vulgar
83	42	43.3	226	2	Q8KF28	Q8KF28 chlorobium
84	42	43.3	238	2	Q835J7	Q835J7 enterococcu
85	42	43.3	244	2	Q8PPV5	Q8PPV5 xanthomonas
86	42	43.3	266	2	Q81D70	Q81D70 bacillus ce
87	42	43.3	296	2	Q7V459	Q7V459 prochloroc
88	42	43.3	315	2	Q6H135	Q6H135 bacillus th
89	42	43.3	319	2	Q9RKM5	Q9RKM5 streptococ
90	42	43.3	379	2	Q8PRC6	Q8PRC6 xanthomonas
91	42	43.3	386	1	ERR1_CANTR	ERR1_CANTR candida tro
92	42	43.3	386	1	ERR2_CANTR	ERR2_CANTR candida tro
93	42	43.3	391	2	Q6U0H2	Q6U0H2 oryza sativ
94	42	43.3	416	1	HEM1_AERPE	Q9Y9J2 aeropyrum p
95	42	43.3	438	2	Q9MFE2	Q9MFE2 beta vulgar
96	42	43.3	443	2	Q89FY2	Q89FY2 bradyrhizob
97	42	43.3	454	2	Q35213	Q35213 oenothera b
98	42	43.3	519	2	Q9SR19	Q9SR19 arabidopsis
99	42	43.3	580	2	Q89RH2	Q89RH2 bradyrhizob
100	42	43.3	581	2	Q43856	Q43856 vicia faba

## ALIGNMENTS

```

RESULT 1
ID 066DR4 PRELIMINARY; PRT; 122 AA.
AC 066DR4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=YPTB0979;
OS Yersinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=1535858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenaï-François V., Souza B., Dacheux D., Elliott J.M.,
RA Darbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH20219.1; -.
KM Hypothetical protein.
SQ SEQUENCE 122 AA; 14215 MW; 7C49F0A1E8BC157 CRC64;

Query Match 50.5%; Score 49; DB 2; Length 122;
Best Local Similarity 53.3%; Pred. No. 4.5;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRH 16
DB 92 INDELRRQKTKH 106

RESULT 2
ID 082C84 PRELIMINARY; PRT; 122 AA.
AC 082C84; Q74WQ2; Q7CK17;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein YPO3137 (Hypothetical protein y1047).
GN OrderedLocNames=YPO794, YPO3137, y1047;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=2147043; PubMed=1156360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Baaham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagers K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1126/STB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

```

```

RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414155; CAC92372.1; -.
DR EMBL; AE013708; NAM84628.1; -.
DR EMBL; AE017129; AAS61059.1; -.
DR PIR; A10380; A10380.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 14215 MW; 7C49F0A1E8BC157 CRC64;

Query Match 50.5%; Score 49; DB 2; Length 122;
Best Local Similarity 53.3%; Pred. No. 4.5;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRH 16
DB 92 INDELRRQKTKH 106

RESULT 3
ID P95613 PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM6;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstrom K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH lyser-type DNA-binding domain.
DR EMBL; Y08963; CAA70157.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_Lyser.
DR InterPro; IPR005119; Lyser_subst.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; Lyser_substrate; 1.
DR PRINTS; PRO00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 49.5%; Score 48; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSR 14
DB 204 KGRSLFQWLSSQ 215

RESULT 4
Y745_HELPJ STANDARD; PRT; 327 AA.
ID Y745_HELPJ
AC Q9ZL58;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DB Hypothetical pseudouridine synthase JHP0682 [EC 4.2.1.70]
GN Pseudouridylylase synthase (Uracil hydrolyase).
OS OrderedLocustNames=JHP0682;
OC Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummillo P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC 5'-phosphate + H(2)O.
CC -1- SIMILARITY: Belongs to the pseudouridine synthase rluA family.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001500; AAD06270.1; -.
DR PIR: B71900; B71900.
DR InterPro: IPR006225; Pseud_synth_RLud.
DR InterPro: IPR006145; Pseudou synth.
DR InterPro: IPR006224; Rlu_synth.
DR InterPro: IPR002942; S4_synth.
DR Pfam: PF00849; Pseudou_synth_2; 1.
DR Pfam: PF01479; S4; 1.
DR ProDom: PD001819; Pseudou_synth; 1.
DR SMART: SM00363; S4; 1.
DR TIGRFAMs: TIGR00005; rluD subfam; 1.
DR PROSITE: PS01129; PSI_RLU; 1.
DR PROSITE: PSS0889; S4; 1.
KW Complete proteome; Hypothetical protein; Lyase; RNA-binding.
FT ACT SITE 136 136 S4 RNA-binding.
FT DOMAIN 12 79 By similarity.
SQ SEQUENCE 327 AA; 37722 MW; 7EDC7F6840D818BD CRC64;

Query Match 48.5%; Score 47; DB 1; Length 327;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TIKPTLRQWLKSRH 16
Db 103 SVKEPTLVDMKSONY 118

RESULT 5
Q750V6 PRELIMINARY; PRT; 727 AA.
AC Q750V6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE AGI167CD.
GN ORFNames=AGI167C;
OS Asbya goesypii (Yeast) (Eremothecium goesypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.

```

```

OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 10895;
RA Lersch A., Brachat S., Voegel S.E., Gaffney T., Philippen P.,
RA Dietrich F.S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB016820; AAS54324.1; -.
DR AGI167C; -.
SQ SEQUENCE 727 AA; 82748 MW; 58A66322705F6767 CRC64;

Query Match 48.5%; Score 47; DB 2; Length 727;
Best Local Similarity 56.2%; Pred. No. 70;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 KGPTLRQWLKSRH 18
Db 92 RGETGRSRRDREHGS 107

RESULT 6
GPD2_MYCPA STANDARD; PRT; 332 AA.
ID GPD2_MYCPA
AC P61744;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycerol-3-phosphate dehydrogenase 2 [NAD(P)+] [EC 1.1.1.94] (NAD(P)H-
DE dependent glycerol-3-phosphate dehydrogenase 2).
GN Name=gp32; OrderedLocustNames=MAP4061c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P)(+) =
CC glycerone phosphate + NAD(P)H.
CC -1- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
CC formation.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the NAD-dependent glycerol-3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE017241; AAS06611.1; -.
DR HAMAP: MF_00394; -; 1.
DR ProDom: PD001278; NAD_G32P_C; 1.
DR PROSITE: PS00957; NAD_G32P; FALSE NEG.
KW Complete proteome; NAD; Oxidoreductase; Phospholipid biosynthesis.
SQ SEQUENCE 332 AA; 35220 MW; B149EF540B313DB3 CRC64;

Query Match 47.9%; Score 46.5; DB 1; Length 332;
Best Local Similarity 49.2%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Qy 3 KGPTLRQWLKSRH 15
Db 29 RGPRTL-QWVRSR 40

RESULT 7
Q6YUC3

```

```

ID Q6YUC3 PRELIMINARY; PRT; 126 AA.
AC Q6YUC3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein B116H04.2 (Hypothetical protein
DE B111C03.14).
GN Name=B116H04.2; Synonyms=B111C03.14;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005871; BAD10677.1; -.
DR EMBL: AP005405; BAD10301.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 126 AA; 12558 MW; F316174EC475BAC6 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 126;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIRGPTLRQWLKSRH 15
Db 66 TAAGPATRRWVVKTRQ 80

RESULT 8
Q6C114 PRELIMINARY; PRT; 365 AA.
ID Q6C114;
AC Q6C114;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Similar to tr|O93968 Candida boidinii Formate dehydrogenase.
GN ORFNames=YAI10F15983g;
OS Varrovia lillopolyltica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Varrovia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvigle A., Talla E.,
RA Goffard N., Frangoul L., Aigle M., Anthonard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckertich J.M., Beyne E., Bleykasten C.,
RA Boismere A., Boyer J., Cattoioco L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreuve F., Hennequin C., Janniaux N., Joyet P., Kachouri R.,
RA Kerres A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straud M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenon-Meyer M., Zivanovic I., Bolocin-Fukuhara M., Thierry A.,
RA Bouchier C., Candron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Soucieu J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CLIB99;
RG Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid
CC dehydrogenase family.
DR EMBL: CR812132; CAG78287.1; -.
DR GO: GO:001616; P-oxidoreductase activity, acting on the CH-O. ., IEA.
DR GO: GO:0006564; P.L-serine biosynthesis; IEA.
DR InterPro: IPR006139; 2-Hacid DH.
DR InterPro: IPR006140; 2-Hacid_DH_C.
DR Pfam: PF00389; 2-Hacid_dh; 1.
DR Pfam: PF02826; 2-Hacid_dh_C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Oxidoreductase.
SQ
SEQUENCE 365 AA; 40172 MW; 8AAC7FE8785139E0 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 365;
Best Local Similarity 72.7%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 LRQWLKSRHHT 17
Db 30 LRQWLKSRHHT 40

RESULT 9
Q82R87 PRELIMINARY; PRT; 402 AA.
ID Q82R87;
AC Q82R87;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transposase.
GN OrderedLocNames=SAV256;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacterlia; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL: AP005021; BAC67965.1; -.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004803; F:transposase activity; IEA.
DR GO: GO:0006133; P:DNA transposition; IEA.
DR InterPro: IPR002559; Transposase_11.
DR Pfam: PF01609; Transposase_11; 1.
KW Complete proteome.
SQ
SEQUENCE 402 AA; 43379 MW; 71EC0A91143451BE CRC64;

Query Match 47.4%; Score 46; DB 2; Length 402;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 2 IRG--PTLRQWLK 13
Db 234 IRG--PTLRQWLK 247

```



## RESULT 10

Q8QUJ6 PRELIMINARY; PRT; 941 AA.  
 ID Q8QUJ6  
 AC Q8QUJ6  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ORF14L.  
 OS Infectious spleen and kidney necrosis virus.  
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae;  
 OC Unclassified Iridoviridae.  
 NC NCB1\_TaxID=180170;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21871810; PubMed=11878882; DOI=10.1006/viro.2001.1208;  
 RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,  
 RA Chan S.M.;  
 RT "Complete genome analysis of the mandarin fish infectious spleen and  
 RT kidney necrosis iridovirus."  
 RL Virology 291:126-139 (2001).  
 DR EMBL: AF371960; AAL98838.1; "-"  
 SQ SEQUENCE 941 AA; 106703 MW; EB66398C7F6CE83 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 941;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHNT 17  
 Db 581 VQGPTLRQWLKSRHNT 596

## RESULT 11

Q980N7 PRELIMINARY; PRT; 186 AA.  
 ID Q980N7  
 AC Q980N7  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Adenylate cyclase, cyab-type, putative (Cyab) (EC 4.6.1.1).  
 GN Name=Cyab; OrderedLocustNames=SS00253;  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 NC NCB1\_TaxID=22287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;  
 RA She O., Singh R.K., Contaloniieri F., Zivancovic Y., Allard G.,  
 RA Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).  
 DR EMBL: AE006661; AAK40592.1; "-"  
 DR PIR: A90167; A90167.  
 DR GO: GO:0004016; F:adenylate cyclase activity; IEA.  
 DR GO: GO:0016829; F:lyase activity; IEA.  
 DR GO: GO:0006171; P:cAMP biosynthesis; IEA.  
 DR InterPro: IPR008172; Adenylate\_cyc.  
 DR InterPro: IPR008173; Cyab.  
 DR Pfam: PF01928; CYTH; 1.  
 DR Pfam: PD009560; Cyab; 1.  
 DR TrEMBL: TIGR00318; cyab; 1.  
 KW Complete proteome; Lyase.

Qy SEQUENCE 186 AA; 21820 MW; 1B47C630B438C868 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 186;  
 Best Local Similarity 55.6%; Pred. No. 33;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQWLKSRHNTS 18  
 Db 65 TYKGPTLRHSLKAREIS 82

## RESULT 12

Q8XR40 PRELIMINARY; PRT; 243 AA.  
 ID Q8XR40  
 AC Q8XR40  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein Rsp1028.  
 GN Name=RS02365; OrderedLocustNames=Rsp1028;  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 NC NCB1\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM1000;  
 RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;  
 RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mengnot S.,  
 RA Ariat M., Billault A., Broctier P., Camus J.C., Catolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Siguer P., Thebaud P., Boucher C.A.,  
 RA Weissenbach J., Bouchet C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502 (2002).  
 DR EMBL: AL646082; CAD18179.1; "-"  
 KW Complete proteome; Hypothetical protein; Plasmid.  
 SQ SEQUENCE 243 AA; 27220 MW; 2B941BEADAF832 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 243;  
 Best Local Similarity 61.5%; Pred. No. 44;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 PTLRQWLKSRHNT 17  
 Db 149 PGLRNLRSRROT 161

## RESULT 13

Q8T462 PRELIMINARY; PRT; 286 AA.  
 ID Q8T462  
 AC Q8T462  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE AT14189p.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Munoz J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiliker S.;  
 RL submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY089335; AAL90073.1; "-"  
 DR FlyBase: FBgn0063731; BCDNA:AT14183.  
 KW Complete proteome; Lyase.

Qy SEQUENCE 286 AA; 30787 MW; 99374B2615D88594 CRC64;

[illegible]

DR	GO:0003723; F:RNA binding; IEA.
DR	GO:0006396; P:RNA processing; IEA.
DR	InterPro; IPR006145; Pseudou_synth.
DR	InterPro; IPR006225; Pseud_synth_RLUD.
DR	InterPro; IPR002942; S4.
DR	Pfam; PF00849; Pseudou_synth_2; 1.
DR	Pfam; PF01479; S4; 1._synth_2; 1.
DR	Prodrom; PD001819; Pseudou_synth; 1.
DR	SMART; SM00363; S4; 1.
DR	TIGRFAMs; TIGR00005; rlud_subfam; 1.
DR	PROSITE; PS50889; S4; 1.
KM	Complete proteome.
SQ	SEQUENCE 351 AA; 40281 MW; 9C5B9C3E7733F6DE CRC64;
Qy	Query Match 46.4%; Score 45; DB 2; Length 351; Best Local Similarity 50.0%; Pred. No. 66; Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Dd	114 SVKEPTLVDMKLNH 129 ::          1 TIKGPTLRWLKSRHH 16
RESULT 16	
O6N3E8	PRELIMINARY; PRT; 789 AA.
AC O6N3E8;	
DT 05-JUL-2004 (TREMBLrel. 27, Created)	
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)	
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)	
DE Possible RND Superfamily transporter.	
CN OrderedLocusNames=RPJ3746;	
OC Rhodopseudomonas palustris;	
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;	
OC Bradyrhizobiaceae; Rhodopseudomonas.	
OX NCBI_TaxID=1076;	
RN [1]	
RP SEQUENCE FROM N.A..	
RC STRAIN=CGA009 / ATCC BAA-98;	
RX PubMed=14704707; DOI=10.1038/nbt923;	
RA Lartner F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,	
RA Land M.L., Pelleterier D.A., Beatty J.T., Lang A.S., Tablica F.R.,	
RA Gibson J.L., Hanson T.E., Bobat C., Torres y Torres J.L., Peres C.,	
RA Harrison F.H., Gibson J., Haywood C.S.;	
RT "Complete genome sequence of the metabolically versatile	
RT photosynthetic bacterium Rhodopseudomonas palustris.";	
RL Nat. Biotechnol. 22:55-61(2004).	
DR EMBL; BX572605; CAE29187.1; --	
DR InterPro; IPR00731; SSD 5TM.	
DR PROSITE; PS50156; SSD; 1.	
KW Complete proteome.	
SQ SEQUENCE 789 AA; 85214 MW; 31C1C9E9CA530CC CRC64;	
Qy	Query Match 46.4%; Score 45; DB 2; Length 789; Best Local Similarity 53.8%; Pred. No. 1.6e+02; Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Dd	6 TLRWMLKSREHTS 18       :   : 523 TLRWMLSEKAHTT 535
RESULT 17	
ID Q7PS78	PRELIMINARY; PRT; 1715 AA.
AC Q7PS78;	
DT 01-MAR-2004 (TREMBLrel. 26, Created)	
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)	
DE ENSANGP0000005643 (Fragment).	
GN Name=ENSANGP00000004323;	
OS Anopheles gambiae str. PESt.	
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	

```

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RC Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
   EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
   preliminary data.
CC EMBL: AAB0100864; EAA06040.2; -.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR InterPro: IPR001395; Aldo/lec. red.
DR InterPro: IPR007735; Pecanex_C.
DR Pfam: PF05041; Pecanex_C.1.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
FT NON TER 1715 1715
SQ SEQUENCE 1715 AA; 189675 MW; ADP294494F2B236A CRC64;

Query Match 46.4%; Score 45; DB 2; Length 1715;
Best Local Similarity 37.5%; Pred. No. 3.9e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

OY 1 TIKGPTLRQWLKSR-----EHTS 18
   ||| ||| ||| ||| |||
ID P93490 PRELIMINARY; PRT; 161 AA.
AC P93490;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cell wall invertase II (Fragment).
OS Pisum sativum (garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=var. final; TISSUE=seed coat;
RA Buchner P.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 32 of glycosyl hydrolases.
CC EMBL: 283339; CAB05954.1; -.
DR PIR: T06826; T06826.
DR GO: GO:0004553; F: hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO: GO:0005975; P: carbohydrate metabolism; IEA.
DR InterPro: IPR001362; Glyco_hydro_32.
DR InterPro: IPR011040; Stalidase.
DR Pfam: PF00251; Glyco_hydro_32; 1.
DR SMART: SM00640; Glyco_32; 1.
DR Glycosidase; Hydrolase.
FT NON TER 1 1
FT NON TER 161 161
SQ SEQUENCE 161 AA; 18033 MW; 32E6B0767F4ABD6 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 161;
Best Local Similarity 46.7%; Pred. No. 41;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKGPTLRQWLKSRH 16
   ||| ||| ||| ||| |||
ID P90433 PRELIMINARY; PRT; 313 AA.
AC P90433;

```

```

DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Truncated reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith J.M., Krauselburd E.N., Torres J.V.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family A2.
CC EMBL: U83413; AAB41428.1; -.
DR HSP: Q07387; ITCW.
DR MEROPS: A02.002; -.
DR GO: GO:0004190; F: aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F: peptidase activity; IEA.
DR GO: GO:0003723; F: RNA binding; IEA.
DR GO: GO:0003964; F: RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F: transferase activity; IEA.
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P: RNA-dependent DNA replication; IEA.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept. Aspartic.
DR InterPro: IPR001969; Pept. Asp_AS.
DR InterPro: IPR00477; RVTse.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT_1; 1.
DR PROSITE: PS00141; ASP_PROT_RETROV; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
RN Transferrase.
FT NON TER 1 1
SQ SEQUENCE 313 AA; 34674 MW; 5A0BB016783FC8A6 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 313;
Best Local Similarity 61.5%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 KGPTLRQWLKSR 15
   ||| ||| ||| ||| |||
ID 184 EGPRLRQWLKSR 196
DB 184 EGPRLRQWLKSR 196

RESULT 20
O8G7M6 PRELIMINARY; PRT; 347 AA.
ID O8G7M6;
AC O8G7M6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE DTP-glucose 4,6-dehydratase enzyme involved in rhamnose
DE biosynthesis.
DE Name=tnlB1; Ordered locus Names=BL0229;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schnell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pesi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
   to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC -1- SIMILARITY: Belongs to the sugar epimerase family.
CC EMBL: AE014641; AAN24075.1; -.
DR HSP: P95780; IKEP.
DR GO: GO:0003824; F: catalytic activity; IEA.

```

DR GO; GO:0008460; F:dtDP-glucose 4,6-dehydratase activity; IEA.  
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.  
 DR InterPro; IPR005888; dtDP\_gluC\_dehyd.  
 DR InterPro; IPR001509; Epimerase\_Dh.  
 DR Pfam; PF01370; Epimerase; 1.  
 DR TIGRfam; TIGR01181; dtDP\_gluC\_dehyd; 1.  
 DR Complete proteome; NAD.  
 KW SEQUENCE 347 AA; 39388 MW; 34852801FD1334FD CRC64;

Query Match 45.4%; Score 44; DB 2; Length 347;  
 Best Local Similarity 37.5%; Pred. No. 96;  
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KGPTRLQWLKSRHTS 18  
 Db 222 KGENVRDWMHTEDHSS 237

RESULT 21  
 Q82PX5 PRELIMINARY; PRT; 377 AA.  
 AC Q82PX5;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=SAV747;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Shimura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinoue M., Takahashi Y., Horikawa H., Nakakawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinoue M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005023; BAC68457.1; -;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 377 AA; 41307 MW; 0253176AAAE62E3 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 377;  
 Best Local Similarity 46.7%; Pred. No. 1e+02;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IKGPTLQWLKSRHTS 16  
 Db 168 MEGPDRLAWLPNRRY 182

RESULT 22  
 O6AKH3 PRELIMINARY; PRT; 500 AA.  
 AC O6AKH3;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Related to penicillin-binding protein (Pbpa).  
 GN OrderedLocustNames=DP2423;

OS Desulfotalea psychrophila.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;  
 OC Desulfobulbaceae; Desulfotalea.  
 OX NCBI\_TaxID=84980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LSv54 / DSM 12343;  
 RX PubMed=15305914;  
 RA Rabus R., Ruepp A., Frickey T., Ralteit T., Fartmann B., Stark M.,  
 RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,  
 RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,  
 RA Klenk H.-P.;

RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium  
 RT from permanently cold Arctic sediments.";  
 RL Environ. Microbiol. 6:887-902(2004).  
 DR EMBL; CR522870; CAG37152.1; -;  
 DR GO; GO:0008658; P:penicillin binding; IEA.  
 DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.  
 DR InterPro; IPR005311; PBP\_dimer.  
 DR InterPro; IPR01460; PBP\_bind\_tpept.  
 DR Pfam; PF03717; PBP\_dimer; 1.  
 DR Pfam; PF00905; Transpeptidase; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 500 AA; 54890 MW; E09B1C5EPD06E554 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 500;  
 Best Local Similarity 47.4%; Pred. No. 1.4e+02;  
 Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 2 IKG-PTLQWLKSRHTS 18  
 Db 92 LKGLKTLNLSWLAGRHSS 110

RESULT 23  
 Q9P888 PRELIMINARY; PRT; 648 AA.  
 AC Q9P888;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative MFS membrane transporter (Fragment).  
 GN Name=mtf;  
 OS Gibberella fujikuroi (Bakane and foot rot disease fungus) (Fusarium  
 OS moniliforme).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella;  
 OC Gibberella fujikuroi complex.  
 OX NCBI\_TaxID=5127;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ms67;  
 RX MEDLINE=21416226; PubMed=11525413;  
 RA Voss T., Schulte J., Tudzyński B.;

RT "A new MFS transporter gene next to the gibberellin biosynthesis gene  
 RT cluster of Gibberella fujikuroi is not involved in gibberellin  
 RT secretion.";  
 RL Curr. Genet. 39:377-383(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 DR EMBL; AJ272424; CAB75959.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005511; F:sugar porter activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0008643; P:carbohydrate transport; IEA.  
 DR InterPro; IPR005828; Sub transporter.  
 DR InterPro; IPR003663; Sugar transprt.  
 DR InterPro; IPR005829; Sug\_transporter.  
 DR Pfam; PF00083; Sugar\_tr; 1.  
 DR PRINTS; PR00171; SUGARTRANSPORT.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; UNKNOWN\_1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Sugar transport; Transmembrane; Transport.

FT NON TER 1 1  
SQ SEQUENCE 648 AA; 72248 MW; 4C90EE9E25AF9C CRC64;

Query March 45.4%; Score 44; DB 2; Length 648;  
Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 QMLKREHT 17  
Db 122 QMLKSOKHT 130

RESULT 24  
Q6VG40 PRELIMINARY; PRT; 1017 AA.  
ID Q6VG40;  
AC Q6VG40;  
DT 05-JUN-2004 (TrEMBLrel. 27, Created)  
DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
DE Pol protein (fragment).  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
NCBI\_TaxID=11723;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=22972551; PubMed=14610175;  
RX DOI=10.1128/JVI.77.23.12523-12534.2003;  
RA Cournaud V., Abela B., Pourrut X., Mpoudi-E Ngole E., Loul S.,  
Delaporte E., Peeters M.;  
RT Identification of a new simian immunodeficiency virus lineage with a  
vpu gene present among different cercopithecus monkeys (C. mona, C.  
cephus, and C. nictitans) from Cameroon.";  
RL J. Virol. 77:12523-12534(2003).  
CC -1- SIMILARITY: Belongs to peptidase family A2.  
EMBL: AY340701; AAR02377.1; -.  
DR HSSP; P12497; 1B9F.  
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0008907; F:integrase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
DR GO; GO:0008270; F:transferase activity; IEA.  
DR GO; GO:0005074; P:DNA integration; IEA.  
DR GO; GO:0006310; P:DNA recombination; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
DR InterPro: IPR01037; Integrase\_C.  
DR InterPro: IPR003308; Peptidase\_A2.  
DR InterPro: IPR001995; Peptidase\_Zn\_N.  
DR InterPro: IPR009007; Pept\_Asp\_Artic.  
DR InterPro: IPR001969; Pept\_Asp\_Artic.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; RNaseH.  
DR InterPro: IPR000477; RNaseH.  
DR InterPro: IPR010659; RVT\_connect.  
DR InterPro: IPR010651; RVT\_thumb.  
DR Pfam; PF02022; Integrase\_Zn\_1.  
DR Pfam; PF00075; RNaseH; 1.  
DR Pfam; PF00665; rve; 1.  
DR Pfam; PF00077; RVP; 1.  
DR Pfam; PF00078; RVT\_1; 1.  
DR Pfam; PF06815; RVT\_connect; 1.  
DR Pfam; PF06817; RVT\_thumb; 1.  
DR PROSITE; PS00141; ASP\_PROTASE; 1.  
DR PROSITE; PS50175; ASP\_PROT\_RETROV; 1.  
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;  
Transferase.  
FT NON TER 1 1  
SQ SEQUENCE 1017 AA; 114690 MW; A1CFE26C001B6B35 CRC64;

Query March 45.4%; Score 44; DB 2; Length 1017;  
Best Local Similarity 61.5%; Pred. No. 3.2e+02;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTRQMLKRE 15  
Db 180 EGPKLKQWPLSRE 192

RESULT 25  
POL\_SIVS4 STANDARD; PRT; 1019 AA.  
ID POL\_SIVS4  
AC P1502;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Pol polyprotein [Contains: Protease (Retropepin) (EC 3.4.23.-);  
Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);  
Integrase (IN)].  
GN Name=POL;  
OS Simian immunodeficiency virus (P236/sm4 isolate) (sooty mangabey).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
NCBI\_TaxID=11737;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=89262053; PubMed=2786147; DOI=10.1038/339389a0;  
RX Hirsch V.M., Olmstead R.A., Murphy-Corb M., Farcell R.H.,  
Johnson P.R.;  
RA "An African primate lentivirus (SIVsm) closely related to HIV-2";  
RL Nature 339:389-392(1989).  
CC -1- FUNCTION: During replicative cycle of retroviruses, the reverse-  
transcribed viral DNA is integrated into the host chromosome by  
the viral integrase enzyme. RNase H activity is associated with  
the reverse transcriptase.  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
phosphomonoester.  
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
+ {DNA}(N).  
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be  
determined.  
CC -1- SIMILARITY: Belongs to the retroviruses pol polyprotein family.  
CC -1- SIMILARITY: Contains 1 integrase-type zinc finger.  
CC -1- SIMILARITY: Contains 1 peptidase A2 domain.  
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.  
CC -1- SIMILARITY: Contains 1 RNase H domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X14307; -; NOT\_ANNOTATED\_CDS.  
DR HSSP; P04584; 1M02.  
DR MEROPS; A02.002; -.  
DR HIV; X14307; POLSSM4.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_Zn\_N.  
DR InterPro: IPR001995; Peptidase\_A2.  
DR InterPro: IPR009007; Pept\_Asp\_Artic.  
DR InterPro: IPR001969; Pept\_Asp\_Artic.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; RNaseH.  
DR InterPro: IPR000477; RNaseH.  
DR InterPro: IPR010659; RVT\_connect.  
DR InterPro: IPR010661; RVT\_thumb.  
DR Pfam; PF00552; Integrase; 1.  
DR Pfam; PF02022; Integrase\_Zn\_1.  
DR Pfam; PF00075; RNaseH; 1.  
DR Pfam; PF00665; rve; 1.  
DR Pfam; PF00077; RVP; 1.



```
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR009007; Pept_Aspartic.
DR InterPro: IPR001969; Pept_Asp_AS.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR010659; RVT_connect.
DR InterPro: IPR010661; RVT_chumb.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF02022; Integrase_Zn; 1.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT_1; 1.
DR Pfam: PF06815; RVT_connect; 1.
DR Pfam: PF06817; RVT_chumb; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; UNKNOWN 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferrase.
FT NON_TER
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKRSRE 15
Db 184 EGPTLRQWLKRSRE 196

RESULT 28
07ZBR7 PRELIMINARY; PRT; 1019 AA.
AC 07ZBR7;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Pol (fragment).
GN Name-pol;
OS Chimpazee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22628501; Pubmed=12743298;
RA DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Deighan H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -1 SIMILARITY: Belongs to peptidase family A2.
DR EMBL; AY21514; AAC67307.1; -.
DR HSSP; P04584; IMU2.
DR GO: GO:0003167; P:DNA binding; IEA.
DR GO: GO:0008907; F:Integrase activity; IEA.
DR GO: GO:0008923; F:peptidase activity; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0015074; P:DNA integration; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn_N.
DR InterPro: IPR001995; Peptidase_A2.
```

```
DR InterPro: IPR009007; Pept_Aspartic.
DR InterPro: IPR001969; Pept_Asp_AS.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR010659; RVT_connect.
DR InterPro: IPR010661; RVT_chumb.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF02022; Integrase_Zn; 1.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT_1; 1.
DR Pfam: PF06815; RVT_connect; 1.
DR Pfam: PF06817; RVT_chumb; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; UNKNOWN 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferrase.
FT NON_TER
SQ SEQUENCE 1019 AA; 115340 MW; A86525DF1BE26F CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 3.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKRSRE 15
Db 184 EGPTLRQWLKRSRE 196

RESULT 29
POL_HV2D2 STANDARD; PRT; 1058 AA.
AC P15833;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropepin) (EC 3.4.23.47);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)].
GN Name-pol;
OS Human immunodeficiency virus type 2 (isolate D205.7) (HIV-2).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11716;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90081881; Pubmed=2594088; DOI=10.1038/342948a0;
RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehnelt H.,
RA Rubsamen-Waigmann H.;
RT "A highly divergent HIV-2-related isolate.";
RL Nature 342:948-950(1989).
CC -1 FUNCTION: During replicative cycle of retroviruses, the reverse-
CC transcribed viral DNA is integrated into the host chromosome by
CC the viral integrase enzyme. RNase H activity is associated with
CC the reverse transcriptase.
CC -1 CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphononucleoside.
CC -1 CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1 CATALYTIC ACTIVITY: Endopeptidase for which the P1 residue is
CC preferably hydrophobic.
CC -1 PFM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1 SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC -1 SIMILARITY: Contains 1 integrase-type zinc finger.
CC -1 SIMILARITY: Contains 1 peptidase A2 domain.
CC -1 SIMILARITY: Contains 1 reverse transcriptase domain.
CC -1 SIMILARITY: Contains 1 RNase H domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```







RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.F.,  
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,  
 RA Hance I., Chris Lee P., Holtzapfe E.K., Scanlan D., Tran K.,  
 RA Moazzaz A., Ullarback T.R., Rizzo M., Lee K., Kosack D., Mestl D.,  
 RA Wedler H., Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Thewissen B.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile *Pseudomonas putida* KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002).  
 DR EMBL; AE016782; AAN67743.1; -.  
 DR HSSP; P03810; 10SA.  
 DR TIGR; PP2130; -.  
 DR InterPro; IPR008939; Muramidase\_bact.  
 DR InterPro; IPR008258; SLT.  
 DR Pfam; PF01464; SLT.1.  
 KM Complete proteome.  
 SQ SEQUENCE 657 AA; 75337 MW; 034030CFD912790 CRC64;

Query Match 44.8%; Score 43.5; DB 2; Length 657;  
 Best Local Similarity 56.2%; Pred. No. 2.4e+02;  
 Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 4 GP-TLRQWKSREHTS 18  
 Db 594 GPGRRQWLKGAHLIS 609

RESULT 33  
 098183 PRELIMINARY; PRT; 63 AA.  
 AC 098183; 012598; 012879;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MC012L (Hypothetical protein B-M.N.L.2).  
 GN Name=MC012L; Synonyms=B-M.N.L.2;  
 OS Molluscum contagiosum virus subtype 1 (MCV1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Molluscipoxvirus.  
 NC NCBI\_Taxid=10280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96325459; PubMed=8670425;  
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,  
 RA Moss B.;  
 RT "Genome sequence of a human tumorigenic poxvirus: prediction of  
 RT specific host response-evasion genes.";  
 RL Science 273:813-816(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97093414; PubMed=8938976;  
 RA Martin-Gallardo A., Moratilla M., Funes J.M., Agromayor M., Nunez A.,  
 RA Varas A.J., Collado M., Valencia A., Lopez-Esteban J.L.,  
 RA Esteban M.;  
 RT "Sequence analysis of a Molluscum contagiosum virus DNA region which  
 RT includes the gene encoding protein kinase 2 and other genes with  
 RT unique organization.";  
 RL Virus Genes 13:19-29(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97352177; PubMed=9208457; DOI=10.1023/A:1007991508159;  
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,  
 RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;  
 RT "A random DNA sequencing, computer-based approach for the generation  
 RT of a gene map of molluscum contagiosum virus.";  
 RL Virus Genes 14:73-80(1997).  
 DR EMBL; U60315; AAC55140.1; -.  
 DR EMBL; U86894; AAB57932.1; -.  
 DR PIR; T30614; T30614.  
 KW Hypothetical protein.  
 SO SEQUENCE 63 AA; 7088 MW; 1C96B36D3E5D8F27 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 63;  
 Best Local Similarity 41.2%; Pred. No. 21;  
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWKSREHTS 18  
 Db 41 VIGETLRATWSRKNTA 57

RESULT 34  
 08VLU9 PRELIMINARY; PRT; 113 AA.  
 AC 08VLU9;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative nitrate reductase (Fragment).  
 GN Name=narg;  
 OS uncultured bacterium.  
 OC Bacteria; environmental samples.  
 OC NCBI\_Taxid=77133;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gregory L.G.;  
 RL Thesis (2000), Department of School of Biological Sciences, University  
 RL of East Anglia, Norwich, United Kingdom.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gregory L.G.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Spiro S.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ314989; CAC85822.1; -.  
 DR HSSP; P09152; 1016.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 FT NON\_TER 1 1  
 FT NON\_TER 113 113  
 SQ SEQUENCE 113 AA; 12691 MW; 34F772F88634988D CRC64;

Query Match 44.3%; Score 43; DB 2; Length 113;  
 Best Local Similarity 35.3%; Pred. No. 40;  
 Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWKSREHTS 18  
 Db 15 IRGVLMLMRKAKEHSA 31

RESULT 35  
 048532 PRELIMINARY; PRT; 124 AA.  
 ID 048532;  
 AC 048532;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Excreted protein (Fragment).  
 GN Name=exca;  
 OS Leptochrix discophora.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Leptochrix.  
 NC NCBI\_Taxid=89;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=SS-1;  
 RA Corstjens P.L.;  
 RL Thesis (1993), Biochemistry, Leiden University.  
 DR EMBL; Z25772; CAA81034.1; -.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR006059; SBP\_bac\_1.  
 DR Pfam; PF01547; SBP\_bac\_1; 1.

```

FT NON TER 1 1
RT NON TER 124 124
SQ SEQUENCE 124 AA; 13393 MW; 33F814295B8BF475 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 124;
Best Local Similarity 54.5%; Pred. No. 44;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLK 12
    |||:::|
Db 40 IKGPSIOEMAK 50

RESULT 36
ID 06AM22 PRELIMINARY; PRT; 215 AA.
AC 06AM22;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Related to hemolysin III.
OS OrderedLocustNames=DP1874;
OC Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSv54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Aman V., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Aman R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG36603.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004254; HlyIII_related.
DR Pfam; PF03006; HlyIII; 1.
KM Complete proteome.
SQ SEQUENCE 215 AA; 23870 MW; F73016495673A459 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 215;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 KGPTLRQWLKSRH 16
    |||:::|
Db 73 KKPVLAKWLRRCDH 86

RESULT 37
ID PSAL_SPTOL STANDARD; PRT; 216 AA.
AC 041385;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Photosystem I reaction center subunit XI, chloroplast precursor (PSI-
DE L) (PSI subunit V).
GN Name=PSAL;
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Monacol.
RX MEDLINE=93344519; PubMed=8343606;
RA Flieger K., Oelmüller R., Herrmann R.G.;

```

```

RT "Isolation and characterization of cDNA clones encoding a 18.8 kDa
RT polypeptide, the product of the gene psal, associated with photosystem
RT I reaction center from spinach.";
RL Plant Mol. Biol. 22:703-709(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
CC thylakoid membrane (Probable).
CC -1- SIMILARITY: Belongs to the psal family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X64445; CAA5775.1; -.
DR PIR; S35151; S35151.
DR HSSP; P25902; 1JBO.
DR InterPro; IPR003757; PSI_Psal.
DR Pfam; PF02605; Psal; 1.
DR ProDom; PD005947; PSI_Psal; 1.
KM Chloroplast; Photosynthesis; Photosystem I; Thylakoid;
KM Transl. peptide; Transmembrane.
FT TRANSIT 1 47 Chloroplast (Potential).
FT CHAIN 48 216 Photosystem I reaction center subunit XI.
FT DOMAIN 48 134 Stromal (Potential).
FT TRANSMEM 135 155 Potential.
FT DOMAIN 156 188 Lumenal (Potential).
FT TRANSMEM 189 209 Potential.
FT DOMAIN 210 216 Stromal (Potential).
SQ SEQUENCE 216 AA; 22937 MW; 603DCA983C7C383B CRC64;

Query Match 44.3%; Score 43; DB 1; Length 216;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRH 18
    |||:::|
Db 27 ISGPALRGPPSPRRH 43

RESULT 38
ID 087115 PRELIMINARY; PRT; 217 AA.
AC 087115;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol. protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SIVagmsABD37;
RX MEDLINE=94298785; PubMed=8026477;
RA Jin M.-J., Hui H., Robertson D.L., Muller M.C., Barre-Sinoussi F.,
RA Hirsch V.M., Allan J.S., Shaw G.M., Sharp P.M., Hahn B.H.;
RT "Mosaic genome structure of simian immunodeficiency virus from west
RT African green monkeys.";
RL EMBL; U04018; AAA21512.1; -.
DR EMBL; U04018; AAA21512.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003664; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Rept. Aspartic.
DR InterPro; IPR000477; RTase.

```

```

DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW RNA-directed DNA polymerase; transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 217 AA; 24503 MW; C1162E4BF18204B8 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 217;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 GPTIRQWLKSR 15
DB 87 GPTIRQWLKSR 98

RESULT 39
O9KC77 PRELIMINARY; PRT; 233 AA.
ID O9KC77;
AC O9KC77;
DT 01-OCT-2000 (TREMBlrel. 15; Created)
DT 01-OCT-2000 (TREMBlrel. 15; Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26; Last annotation update)
DE Initiation of chromosome replication.
GN Name=dnad; OrderedLocustNames=BH1697;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20515582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001512; BAB05416.1; -.
DR EMBL; AP001512; BAB05416.1; -.
PTR: A83862; A83862.
DR InterPro; IPR006343; Dnad_phage.
DR InterPro; IPR009058; Wing_Nlx_DNA_bnd.
DR Pfam; PF04271; Dnad; 1.
DR TIGRFAMs; TIGR01446; Dnad_dom; 1.
KM Complete proteome.
SQ SEQUENCE 233 AA; 27045 MW; D25E82126AD97CC3 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 233;
Best Local Similarity 43.8%; Pred. No. 89;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 IKGPTIRQWLKSRHT 17
DB 144 IEGTILSMWIDQDHT 159

RESULT 40
PYRF_GLOVI STANDARD; PRT; 237 AA.
ID PYRF_GLOVI;
AC Q7NK22;
DT 29-MAR-2004 (Rel. 43; Created)
DT 29-MAR-2004 (Rel. 43; Last sequence update)
DT 05-JUN-2004 (Rel. 44; Last annotation update)
DE Oxidative 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPdecase) (OMPdecase).
GN Name=pyrf; OrderedLocustNames=g11658;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;

```

```

RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Saenamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003)
CC -1- CATALYTIC ACTIVITY: Oxidative 5'-phosphate = UMP + CO(2).
CC -1- PYRHWAY: Pyrimidine biosynthesis, sixth (last) step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: Belongs to the OMP decarboxylase family. Subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AP006573; BAC89599.1; -.
DR HAMAP; MF_01200; -.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR TIGRFAMs; TIGR01740; pyrf; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KM Complete proteome; Decarboxylase, lyase; Pyrimidine biosynthesis.
FT ACT SITE 62
FT ACT SITE 62 Proton donor (By similarity).
SQ SEQUENCE 237 AA; 24470 MW; 0959AC5628E8258C CRC64;

Query Match 44.3%; Score 43; DB 1; Length 237;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 TIKGPTIRQWLKSR 14
DB 40 TIAQPGIIEWLKAR 53

RESULT 41
O9XSK8 PRELIMINARY; PRT; 238 AA.
ID O9XSK8;
AC O9XSK8;
DT 01-MAR-2002 (TREMBlrel. 20; Created)
DT 01-MAR-2002 (TREMBlrel. 20; Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24; Last annotation update)
DE Hypothetical protein RSP0463.
GN Name=RS00951; OrderedLocustNames=RS0463;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
RA Ariat M., Billault A., Broctier P., Camus J.C., Cactolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunne S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiek T.,
RA Sigler P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646078; CAD17614.1; -.
KM Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 238 AA; 25530 MW; ABA94D28568858E7 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 238;
Best Local Similarity 61.5%; Pred. No. 92;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

OY 5 PTIRWLKSRHT 17  
 DB 221 PIRNGWLKRLHT 233

RESULT 42  
 ID YMO8 PARTE STANDARD; PRT; 241 AA.  
 AC P15609;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Hypoetical 28.6 kDa protein (ORF8).  
 OS Paramecium tetraurelia.  
 OG Mitochondrion.  
 OC Eukaryote; Alveolata; Ciliophora; Oligohymenophorea; Periculida;  
 OC Paramecium.  
 OX NCBI\_TaxID=5888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Strock 51;  
 RX MEDLINE=90174913; PubMed=2308823;  
 RA Pritchard A.E., Sellhauer J.U., Mahalingam R., Sable C.L.,  
 RA Venuti S.B., Cummings D.U.;  
 RT "Nucleotide sequence of the mitochondrial genome of Paramecium";  
 RL Nucleic Acids Res. 18:173-180(1990).  
 CC -1- SIMILARITY: Belongs to the ribosomal protein S13p family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X15917; CAA34049.1; -.  
 DR PIR; S07740; S07740.  
 DR InterPro; IPR010979; Ribosomal\_H2TH.  
 DR InterPro; IPR001892; Ribosomal\_S13.  
 DR Pfam; PF00416; Ribosomal\_S13\_1.  
 DR PROSITE; PS00646; RIBOSOMAL\_S13\_1; FALSE\_NEG.  
 DR PROSITE; PS50159; RIBOSOMAL\_S13\_2; 1.  
 KM Hypothetical protein; Mitochondrion.  
 SQ SEQUENCE 241 AA; 26648 MW; 7410BA96B37FA6F CRC64;  
 Query Match 44.3%; Score 43; DB 1; Length 241;  
 Best Local Similarity 41.2%; Pred. No. 93;  
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
 OY 2 IKGPTLRLKSRHTS 18  
 DB 26 VKGPTLRLKRFPRYNA 42

RESULT 43  
 ID RIBF MYCPN STANDARD; PRT; 269 AA.  
 AC P75587;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Putative riboflavin biosynthesis protein ribf [includes: Riboflavin  
 DE kinase (EC 2.7.7.2) (FAD pyrophosphorylase), FMN adenylyltransferase  
 DE (EC 2.7.7.2) (FAD pyrophosphorylase), FMN synthetase]].  
 GN Name=ribf; OrderedLocNames=MPN158; ORFNames=mp673;  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;  
 RA Himmelfreuch R., Hilbert H., Plegens H., Pirkl E., Li B.-C.,  
 RA Hermann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae";  
 RT Nucleic Acids Res. 24:4420-4449(1996).  
 CC -1- CATALYTIC ACTIVITY: ATP + riboflavin = ADP + FMN.  
 CC -1- CATALYTIC ACTIVITY: ATP + FMN = diphosphate + FAD.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE000062; AAB96321.1; -.  
 DR PIR; S73999; S73999.  
 DR HSSP; Q96966; INB9.  
 DR InterPro; IPR002606; FAD\_Synth.  
 DR Pfam; PF01687; FAD\_Synth; 1.  
 DR Pfam; PF06574; Flavokinase; 1.  
 DR Prodom; PD003662; FAD\_Synth; 1.  
 DR TIGRfam; TIGR00083; Ribf; 1.  
 KM Complete proteome; FAD; FMN; Multifunctional enzyme;  
 KM Nucleotidyltransferase; Transferase.  
 SQ SEQUENCE 269 AA; 30435 MW; 2E63D7BC7A8FA12D CRC64;  
 Query Match 44.3%; Score 43; DB 1; Length 269;  
 Best Local Similarity 58.3%; Pred. No. 1e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 TIKGPTLRLK 12  
 DB 124 TLSSSTIRWLK 135

RESULT 44  
 ID Q8UN03 PRELIMINARY; PRT; 340 AA.  
 AC Q8UN03;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Reverse transcriptase (Fragment).  
 GN Name=pol;  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF458220; AAL59620.1; -.  
 DR HSSP; P04584; 1MU2.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
 DR Pfam; PF00078; RVT\_1; 1.  
 DR Pfam; PF06815; RVT\_connect; 1.  
 DR Pfam; PF06817; RVT\_chumb; 1.  
 KM RNA-directed DNA polymerase; Transferase.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 340 AA; 39547 MW; 7777BFD2A057EA6B CRC64;  
 Query Match 44.3%; Score 43; DB 2; Length 340;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 OY 4 GPTLRLKSRRE 15

Db 18 GPKLRQWPLSKR 29

## RESULT 45

OBUN04 PRELIMINARY; PRT; 340 AA.  
 AC OBUN04;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Reverse transcriptase (Fragment).  
 GN Name-Pol;  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RA Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;  
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF458219; AAL59619.1; -.  
 DR HSSP; P04584; IMU2.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.  
 DR Pfam; PF00078; RVT\_1; 1.  
 DR Pfam; PF06815; RVT\_connect; 1.  
 DR Pfam; PF06817; RVT\_chumb; 1.  
 DR RNA-directed DNA polymerase; Transferase.  
 FT NON\_TER 1  
 FT 340  
 SQ SEQUENCE 340 AA; 39545 MM; F9F3BPD3F4005252 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 340;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GPTLRQWPLSKR 15  
 DB 18 GPKLRQWPLSKR 29

Search completed: September 1, 2005, 16:21:06  
 Job time : 70.9496 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 82.7482 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-10

Perfect score: 98

Sequence: 1 SIBGPTLRWLTSTRTPHS 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database: A\_Geneseq\_16dec04:\*

1: geneseqp19808:\*\n2: geneseqp19908:\*\n3: geneseqp20008:\*\n4: geneseqp20018:\*\n5: geneseqp20028:\*\n6: geneseqp20038:\*\n7: geneseqp20048:\*\n8: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	18	2	AAW09460 Thrombopo
2	98	100.0	18	2	AAW09498 Thrombopo
3	98	100.0	18	2	AAW36649 Thrombopo
4	98	100.0	18	2	AAW33027 Thrombopo
5	98	100.0	18	2	AAW36652 Thrombopo
6	98	100.0	18	3	AAW17026 Thrombopo
7	98	100.0	18	4	AAW25868 Human chr
8	98	100.0	18	4	AAU25824 Human chr
9	98	100.0	18	4	AAU25871 Human chr
10	98	100.0	18	5	ABW72912 TPO mimet
11	98	100.0	18	7	ADJ73064 TPO mimet
12	98	100.0	18	8	ADJ52699 CHI delet
13	98	100.0	18	8	ADJ51660 CHI delet
14	98	100.0	18	8	ADJ51693 TPO mimet
15	98	100.0	18	8	ADJ51693 TPO mimet
16	98	100.0	18	8	ADJ51693 TPO mimet
17	98	100.0	18	8	ADJ51693 TPO mimet
18	98	100.0	18	8	ADJ51693 TPO mimet
19	98	100.0	18	8	ADJ51693 TPO mimet
20	98	100.0	18	8	ADJ51693 TPO mimet
21	98	100.0	18	8	ADJ51693 TPO mimet
22	98	100.0	18	8	ADJ51693 TPO mimet
23	98	100.0	18	8	ADJ51693 TPO mimet
24	98	100.0	18	8	ADJ51693 TPO mimet
25	98	100.0	18	8	ADJ51693 TPO mimet

26	66	67.3	18	5	ABP51684 TPO mimet
27	66	67.3	18	5	ABP51690 TPO mimet
28	66	67.3	18	5	ABP51675 TPO mimet
29	66	67.3	18	8	ADQ16611 TPO mimet
30	66	67.3	18	8	ADQ16619 TPO mimet
31	66	67.3	18	8	ADQ16621 TPO mimet
32	66	67.3	18	8	ADQ16646 TPO mimet
33	66	67.3	18	8	ADQ16615 TPO mimet
34	66	67.3	18	8	ADQ16617 TPO mimet
35	66	67.3	18	8	ADQ16623 TPO mimet
36	66	67.3	18	8	ADQ16708 TPO mimet
37	66	67.3	18	8	ADQ16710 TPO mimet
38	66	67.3	128	8	ADQ16705 Modified
39	66	67.3	225	5	ADQ16704 Modified
40	66	67.3	472	8	ADQ16647 TPO mimet
41	66	67.3	472	8	ADQ16647 TPO mimet
42	66	67.3	14	5	ABW16969 TPO mimet
43	66	67.3	14	5	ABW16969 TPO mimet
44	66	67.3	14	5	ABW16969 TPO mimet
45	66	67.3	14	5	ABW16969 TPO mimet
46	66	67.3	14	5	ABW16969 TPO mimet
47	66	67.3	14	5	ABW16969 TPO mimet
48	66	67.3	14	5	ABW16969 TPO mimet
49	66	67.3	14	5	ABW16969 TPO mimet
50	66	67.3	14	5	ABW16969 TPO mimet
51	66	67.3	14	5	ABW16969 TPO mimet
52	66	67.3	14	5	ABW16969 TPO mimet
53	66	67.3	14	5	ABW16969 TPO mimet
54	66	67.3	14	5	ABW16969 TPO mimet
55	66	67.3	14	5	ABW16969 TPO mimet
56	66	67.3	14	5	ABW16969 TPO mimet
57	66	67.3	14	5	ABW16969 TPO mimet
58	66	67.3	14	5	ABW16969 TPO mimet
59	66	67.3	14	5	ABW16969 TPO mimet
60	66	67.3	14	5	ABW16969 TPO mimet
61	66	67.3	14	5	ABW16969 TPO mimet
62	66	67.3	14	5	ABW16969 TPO mimet
63	66	67.3	14	5	ABW16969 TPO mimet
64	66	67.3	14	5	ABW16969 TPO mimet
65	66	67.3	14	5	ABW16969 TPO mimet
66	66	67.3	14	5	ABW16969 TPO mimet
67	66	67.3	14	5	ABW16969 TPO mimet
68	66	67.3	14	5	ABW16969 TPO mimet
69	66	67.3	14	5	ABW16969 TPO mimet
70	66	67.3	14	5	ABW16969 TPO mimet
71	66	67.3	14	5	ABW16969 TPO mimet
72	66	67.3	14	5	ABW16969 TPO mimet
73	66	67.3	14	5	ABW16969 TPO mimet
74	66	67.3	14	5	ABW16969 TPO mimet
75	66	67.3	14	5	ABW16969 TPO mimet
76	66	67.3	14	5	ABW16969 TPO mimet
77	66	67.3	14	5	ABW16969 TPO mimet
78	66	67.3	14	5	ABW16969 TPO mimet
79	66	67.3	14	5	ABW16969 TPO mimet
80	66	67.3	14	5	ABW16969 TPO mimet
81	66	67.3	14	5	ABW16969 TPO mimet
82	66	67.3	14	5	ABW16969 TPO mimet
83	66	67.3	14	5	ABW16969 TPO mimet
84	66	67.3	14	5	ABW16969 TPO mimet
85	66	67.3	14	5	ABW16969 TPO mimet
86	66	67.3	14	5	ABW16969 TPO mimet
87	66	67.3	14	5	ABW16969 TPO mimet
88	66	67.3	14	5	ABW16969 TPO mimet
89	66	67.3	14	5	ABW16969 TPO mimet
90	66	67.3	14	5	ABW16969 TPO mimet
91	66	67.3	14	5	ABW16969 TPO mimet
92	66	67.3	14	5	ABW16969 TPO mimet
93	66	67.3	14	5	ABW16969 TPO mimet
94	66	67.3	14	5	ABW16969 TPO mimet
95	66	67.3	14	5	ABW16969 TPO mimet
96	66	67.3	14	5	ABW16969 TPO mimet
97	66	67.3	14	5	ABW16969 TPO mimet
98	66	67.3	14	5	ABW16969 TPO mimet





Db 1 SIEGPTLRWLTSTRPHS 18

## RESULT 3

ID AAM36649 standard; peptide; 18 AA.

AC AAM36649;

DT 11-MAR-1998 (first entry)

DB Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
 XX haematological disorder; thrombocytopenia; chemotherapy;  
 XX radiation therapy; bone marrow transfusion; diagnosis;  
 XX signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matcheak's LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 27; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.1e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIEGPTLRWLTSTRPHS 18

1 SIEGPTLRWLTSTRPHS 18

1 SIEGPTLRWLTSTRPHS 18

RESULT 4  
 ID AAM33027 standard; peptide; 18 AA.

AC AAM33027;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matcheak's LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a  
 CC molecular weight of less than 8000 Da and a TR binding affinity as  
 CC expressed by an IC50 of no more than about 100 microm. It can be used to  
 CC treat disorders which are susceptible to treatment with a thrombopoietin  
 CC agonist, preferably haematological disorders and thrombocytopenia  
 CC resulting from chemotherapy, radiation therapy or bone marrow  
 CC transfusions. It can also be used diagnostically, e.g. to investigate the  
 CC mechanism of thrombopoietin signal transduction and receptor activation,  
 CC or to maintain the proliferation and growth of thrombopoietin dependent  
 CC cell lines

SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.1e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIEGPTLRWLTSTRPHS 18

1 SIEGPTLRWLTSTRPHS 18

1 SIEGPTLRWLTSTRPHS 18

RESULT 5  
 ID AAM36652 standard; peptide; 18 AA.

AC AAM36652;

DT 11-MAR-1998 (first entry)

DB Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
 XX haematological disorder; thrombocytopenia; chemotherapy;  
 XX radiation therapy; bone marrow transfusion; diagnosis;  
 XX signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX PA  
 PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Matcheakis LC, Schatz PJ, Wagstrom CR, Wighton NC;  
 XX WPI, 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Disclosure; Page 27; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transplants. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 18 AA;  
 XX  
 Query Match 100.0%; Score 98; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SIEGPTLREWLTSRTPHS 18  
 DB 1 SIEGPTLREWLTSRTPHS 18  
 XX  
 RESULT 6  
 AAB17026  
 ID AAB17026 standard; peptide; 18 AA.  
 XX  
 AC AAB17026;  
 XX  
 DT 31-OCT-2000 (first entry)  
 XX  
 DE TPO-mimetic peptide sequence SEQ ID NO:82.  
 XX  
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTAA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200024782-A2.  
 XX  
 PD 04-MAY-2000.  
 XX  
 PF 25-OCT-1999; 99WO-US025044.  
 XX  
 PR 23-OCT-1998; 98US-0105371P.  
 PR 22-OCT-1999; 99US-00428082.  
 XX  
 PA (AMGR-) AMGEN INC.  
 XX  
 PI Feige U, Liu C, Cheetham J, Boone TC;  
 XX  
 DR WPI; 2000-350702/30.  
 XX  
 PT Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 PS Claim 19; Page 222; 608pp; English.  
 XX  
 CC The present invention describes composition of matter (1) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (1) is:  
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-C-P1, -(L1)-C-P1-(L2)-d-P2, -(L1)-C-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;  
 XX  
 Query Match 100.0%; Score 98; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SIEGPTLREWLTSRTPHS 18  
 DB 1 SIEGPTLREWLTSRTPHS 18  
 XX  
 RESULT 7  
 AAU25868  
 ID AAU25868 standard; peptide; 18 AA.  
 XX  
 AC AAU25868;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #54.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Depreince RB, Podduturi S;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure, Col 20, 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 18 AA:

Query Match 100.0%; Score 98; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGPTLRWLTSTRPHS 18  
 |||  
 DB 1 STEGPTLRWLTSTRPHS 18

# RESULT 8

AAU25824  
 ID AAU25824 standard; peptide; 18 AA.

AC AAU25824;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #10.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwila SE, Gates CM, Schatz PJ,  
 PI Balaubramanian P, Wagstrom CR, Hendren RW, Depirnce RB, Poddaturi S,  
 PI Yin Q.

XX WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure, Col 67-68; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 18 AA:

Query Match 100.0%; Score 98; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGPTLRWLTSTRPHS 18  
 |||  
 DB 1 STEGPTLRWLTSTRPHS 18

# RESULT 9

AAU25871  
 ID AAU25871 standard; peptide; 18 AA.

AC AAU25871;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #57.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwila SE, Gates CM, Schatz PJ,  
 PI Balaubramanian P, Wagstrom CR, Hendren RW, Depirnce RB, Poddaturi S,  
 PI Yin Q.

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprising contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 XX  
 CC Sequences AMU5815-AMU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 XX  
 SQ Sequence 18 AA;  
 XX  
 QY Query Match 100.0%; Score 98; DB 4; Length 18;  
 XX Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 1 SIEGPTLRWLTSTRTPHS 18  
 1 SIEGPTLRWLTSTRTPHS 18  
 XX  
 RESULT 10  
 ID ABB72912 standard; peptide, 18 AA.  
 XX  
 AC ABB72912;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:82.  
 XX  
 KW Modified peptide; mimetic; Fe domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antineumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianemic; anorectic; antiinfectivity; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200183525-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 02-MAY-2001; 2001WO-US014310.  
 XX  
 PR 03-MAY-2000; 2000US-00563286.  
 XX  
 PA (AMGB-) AMGEN INC.  
 XX

PI Feige U, Liu C, Cheatham JC, Boone TC, Gudas JM;  
 XX  
 DR WPI; 2002-130313/17.  
 XX  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 44; 176pp; English.  
 XX  
 CC The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antineumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antiinfectivity, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL5695 to ABL5777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;  
 XX  
 QY Query Match 100.0%; Score 98; DB 5; Length 18;  
 XX Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 1 SIEGPTLRWLTSTRTPHS 18  
 1 SIEGPTLRWLTSTRTPHS 18  
 XX  
 RESULT 11  
 ID ADJ73064 standard; peptide, 18 AA.  
 XX  
 AC ADJ73064;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE TPO mimetic peptide sequence SeqID 518.  
 XX  
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KW immunomodulator; cardiac; antimicrobial; cytostatic; neuroprotective;  
 KW TPO.  
 XX  
 XX Synthetic.  
 OS  
 XX  
 PN WO2003084477-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 24-MAR-2003; 2003WO-US009139.  
 XX  
 PR 29-MAR-2002; 2002US-0368791P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Scallion BJ, Grayeb J;  
 XX

DR WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain  
PT variable region comprising human framework or ligand binding region,  
PT useful for preparing a composition for treating e.g., immune,  
PT cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 518; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific  
CC portions or variants thereof. Specifically, it refers to an antibody  
CC fragment where a protein has been inserted into, or replaces a portion  
CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
CC least one portion of a heavy chain or light chain variable region, which  
CC itself comprises at least one human framework region and at least one  
CC ligand binding region (LBR). The present invention describes human  
CC mimetibodies, including modified immunoglobulins and cleavage products  
CC that can be useful in gene therapy and the generation of transgenic  
CC plants and animals. Furthermore, the CDR mimetibody is useful for  
CC preparing compositions for modulating, treating or reducing the symptoms  
CC of immune, cardiovascular, infectious, malignant and/or neurologic  
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
CC peptide sequence is a TPO mimetic peptide sequence used to make a  
CC mimetibody of the invention.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 7; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIEGPTLRRLMTSRTPHS 18  
1 SIEGPTLRRLMTSRTPHS 18  
DB 1 SIEGPTLRRLMTSRTPHS 18

RESULT 12  
ADJ52699

ID ADJ52699 standard; peptide; 18 AA.

XX AC ADJ52699;

XX DT 06-MAY-2004 (first entry)

XX DB CH1 deleted mimetibody-related peptide SeqID518.

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
KW arrhythmia; hypertension; heart failure; neurodegenerative;  
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
KW cancerous condition; infectious disease; bacterial infection;  
KW viral infection; fungal infection.

KW Unidentified.

OS Synthetic.

XX PN MO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (GEN2 ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

XX PI Kutolowski KA;

XX DR WPI; 2004-082870/08.

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
PT disease.

XX PS Claim 2; SEQ ID NO 518; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an immunosuppressive,  
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
CC is useful for diagnosing or treating a disease condition in a cell,  
CC tissue, organ or animal, specifically for modulating, treating,  
CC alleviating, preventing the incidence or reducing the symptoms of an  
CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
CC conditions, or infectious diseases (for example bacterial, viral or  
CC fungal infection). The present sequence is that of a peptide which may be  
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 98; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIEGPTLRRLMTSRTPHS 18  
1 SIEGPTLRRLMTSRTPHS 18  
DB 1 SIEGPTLRRLMTSRTPHS 18

RESULT 13  
ADJ51660

ID ADJ51660 standard; peptide; 18 AA.

XX AC ADJ51660;

XX DT 06-MAY-2004 (first entry)

XX DB CH1 deleted mimetibody-related peptide SeqID518.

KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; tumour necrosis factor;  
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
KW obstetric disorder; haematological disorder; immunological disorder;  
KW allergic disorder; infectious disorder; musculoskeletal disorder;  
KW oncological disorder; neurological disorder; nutritional disorder;  
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
KW renal disorder; pulmonary disorder.

OS Unidentified.

XX PN MO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (GEN2 ) CENTOCOR INC.

PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neseppor TC;  
PI Kutoolski KA;  
XX WPI; 2004-082872/08.  
XX  
XX  
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
PT diagnosing, preventing or treating cardiovascular, dermatologic,  
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
PT nutritional disorders.  
XX  
XX  
XX Claim 15; SEQ ID NO 518; 123pp; English.  
XX  
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences  
XX which encode them), compositions, methods and uses. The invention may be  
XX useful for the development of compounds with an osteopathic,  
XX cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
XX gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
XX immunomodulator, antiallergic, muscular-Gen, cytostatic,  
XX antiinflammatory, neuroleptic, ophthalmological, nephrotropic or  
XX respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
XX modulator or cytokine-agonist. The methods and compositions of the  
XX present invention are useful for the diagnosis, prevention and/or  
XX treatment of diseases or conditions associated with aberrant expression  
XX or activity of the CHI deleted mimetibody, such as a bone or joint,  
XX cardiovascular, dental or oral, dermatological, ear, nose or throat,  
XX endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
XX obstructive, haematologic, immunological, allergic, infectious,  
XX musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
XX pediatric, psychiatric, renal or pulmonary disorders. The present  
XX sequence is that of a peptide which may be used during the creation of a  
XX mimetibody of the invention.  
XX  
XX  
XX Sequence 18 AA;  
SQ  
Query Match 100.0%; Score 98; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SIEGPTLRWMLTSRTPHS 18  
DB 1 SIEGPTLRWMLTSRTPHS 18  
RESULT 14  
ABP51693  
ID ABP51693 standard; peptide; 18 AA.  
XX  
XX ABP51693;  
AC  
XX  
XX  
XX 01-OCT-2002 (first entry)  
DT  
XX  
XX TPO mimetic peptide SEQ ID NO:49.  
DE  
XX  
XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
XX complementarity determining region; immunoglobulin; antianaemic;  
XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
XX  
XX Homo sapiens.  
OS  
XX Synthetic.  
XX  
XX WO200246238-A2.  
XX  
XX  
XX 13-JUN-2002.  
XX  
XX  
XX 05-DEC-2001; 2001WO-US047656.  
XX  
XX  
XX 05-DEC-2000; 2000US-0251448P.  
XX  
XX 04-MAY-2001; 2001US-0288889P.  
XX  
XX 29-MAY-2001; 2001US-0294068P.  
XX  
XX (ALEX-) ALEXION PHARM INC.  
XX  
XX Bowditch KS, Barbas-Frederickson S, Renshaw M;

XX  
XX WPI; 2002-566610/60.  
DR N-PSDB; ABQ73371.  
XX  
XX  
XX A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.  
XX  
XX  
XX Claim 20; Fig 5; 113pp; English.  
XX  
XX  
XX The present invention describes an immunoglobulin molecule or its fragment  
XX (I) comprising a region where amino acid residues corresponding to at  
XX least a portion of the complementary determining region (CDR) are  
XX replaced or fused with biologically active peptides e.g. a peptide  
XX mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
XX that is flanked with proline at its carboxy terminus. (I) has  
XX antianaemic, haemostatic and nephrotropic activities, and can be used as  
XX a stimulator of proliferation, differentiation and maturation of  
XX haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
XX for stimulating proliferation, differentiation or growth of  
XX promegakaryocytes or megakaryocytes, where (I) is contacted with  
XX promegakaryocytes or megakaryocytes, which results in increased platelet  
XX production. (I) with a region where amino acid residues corresponding to  
XX a portion of CDR is replaced with an EPO mimetic, or which has one or  
XX more of its CDRs fused to an EPO mimetic, is useful for increasing the  
XX production of red blood cells, where (I) is contacted with haematopoietic  
XX stem cells or their progenitors. (I) is useful for diagnostics or  
XX therapeutics, in cell isolation strategies, and for treating patients  
XX suffering from deficiency in cell populations caused by disease,  
XX disorders or treatments related to the suppression of haematopoiesis.  
XX ABQ73288 to ABQ73377 and ABP51696 to ABP51696 represent sequences used in  
XX the exemplification of the present invention  
XX  
XX  
XX Sequence 18 AA;  
SQ  
Query Match 68.4%; Score 67; DB 5; Length 18;  
Best Local Similarity 68.8%; Pred. No. 0.0011;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 1 SIEGPTLRWMLTSRTP 16  
DB 2 TIEGPTLRWMLAARAP 17  
RESULT 15  
ABP51691  
ID ABP51691 standard; peptide; 18 AA.  
XX  
XX ABP51691;  
AC  
XX  
XX  
XX 01-OCT-2002 (first entry)  
DT  
XX  
XX TPO mimetic peptide SEQ ID NO:45.  
DE  
XX  
XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
XX complementarity determining region; immunoglobulin; antianaemic;  
XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
XX  
XX Homo sapiens.  
OS  
XX Synthetic.  
XX  
XX WO200246238-A2.  
XX  
XX  
XX 13-JUN-2002.  
XX  
XX  
XX 05-DEC-2001; 2001WO-US047656.  
XX  
XX  
XX 05-DEC-2000; 2000US-0251448P.  
XX  
XX 04-MAY-2001; 2001US-0288889P.  
XX  
XX 29-MAY-2001; 2001US-0294068P.  
XX  
XX (ALEX-) ALEXION PHARM INC.  
XX  
XX

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;  
 PI WPI; 2002-566610/60.  
 XX DR N-PSDB; ABQ73369.  
 XX

PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.  
 XX

PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antanaemic, haemostatic and nephrotoxic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51659 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 CC  
 XX Sequence 18 AA;  
 SQ

Query Match 68.4%; Score 67; DB 5; Length 18;  
 Best Local Similarity 68.8%; Pred. No. 0.0011;  
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SIEGPTLRWLTSTRTP 16  
 :|||||:|:|:  
 DB 2 TIEGPTLRQWLARAP 17

RESULT 16

ID ADQ16625 standard; peptide; 18 AA.

XX ADQ16625;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide with random flanking residues SEQ ID NO:45.

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;

KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;

KW immunotherapy; thrombocytopenia.

XX Unidentified.

XX WO2004050017-A2.

XX 17-JUN-2004.

XX 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.  
 XX DR N-PSDB; ADQ16626.  
 DR

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX

PS Example 1; SEQ ID NO 45; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.  
 XX

SQ Sequence 18 AA;

Query Match 68.4%; Score 67; DB 8; Length 18;  
 Best Local Similarity 68.8%; Pred. No. 0.0011;  
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SIEGPTLRWLTSTRTP 16  
 :|||||:|:|:  
 DB 2 TIEGPTLRQWLARAP 17

RESULT 17

ID ADQ16629 standard; peptide; 18 AA.

XX ADQ16629;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide with random flanking residues SEQ ID NO:49.

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;

KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;

KW immunotherapy; thrombocytopenia.

XX Unidentified.

XX WO2004050017-A2.

XX 17-JUN-2004.

XX 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;

XX WPI; 2004-460973/43.

XX N-PSDB; ADQ16630.

XX New immunoglobulin molecule comprising a region, where two

XX complementarity determining regions (CDRs) are replaced with EPO mimetic

XX or a TPO mimetic, useful for treating thrombocytopenia.  
 XX Example 1; SEQ ID NO 49; 107pp; English.  
 XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a



CC portion of a two complementarily determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

XX Sequence 18 AA;

Query Match 68.4%; Score 67; DB 8; Length 18;

Best Local Similarity 68.8%; Pred. No. 0.0011; Mismatches 2; Indels 0; Gaps 0;

Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SIEGPTLRWMTSRTP 16  
 :|||||:|:|  
 Db 2 TIEGPTLRQWLAAAP 17

RESULT 18

ADN59830  
 ID ADN59830 standard; peptide; 22 AA.

AC ADN59830;

DT 01-JUL-2004 (first entry)

DE TMP peptide TMP12.

XX Haemostatic; antihaemic; immunosuppressive; platelet;

KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;

KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KW autoimmune haemolytic anaemia; Hughes' syndrome;

XX lupoid thrombocytopenia; linker.

OS Homo sapiens.

XX WO2003031589-A2.

XX 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

PA (AMGE-) AMGEN INC.

PI Min H, Stiney KC, Hartley C;

DR WPI; 2003-403101/38.

XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

PT which stimulate the production of platelets and/or the production of

DR platelet precursors, useful for treating thrombocytopenia.

XX Example 6; Page 83; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
 CC platelets in a patient. The TMP of the invention is useful for treating  
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
 CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
 CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
 CC autoimmune haemolytic anaemia, Hughes' syndrome and lupoid

CC thrombocytopenia. The TMP of the invention is also useful for  
 CC maintaining the viability or storage life of platelets and/or  
 CC megakaryocytes and its derived cells. The compounds demonstrate an  
 CC improved ability to bind to and/or trigger transmembrane signal through,  
 CC i.e. activating, the mpl receptor the compounds have superior.  
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
 CC vitro, the production of platelets and/or megakaryocytic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life,  
 CC biological activity and in vivo circulation time. The current sequence  
 CC represents a TMP peptide of the invention to which a two amino acid "cap"  
 CC has been added to the carboxy terminal to increase peptide affinity.

XX Sequence 22 AA;

Query Match 67.9%; Score 66.5; DB 7; Length 22;

Best Local Similarity 72.2%; Pred. No. 0.0017; Mismatches 3; Indels 1; Gaps 1;

Matches 13; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 IEGPTLRWMTSR-TPHS 18  
 :|||||:|:|  
 Db 5 IEGPTLRQWLAAAPLPHS 22

RESULT 19

ADN59708  
 ID ADN59708 standard; peptide; 25 AA.

AC ADN59708;

DT 01-JUL-2004 (first entry)

DE Thrombopoietin mimetic peptide TMP12, seq id 57.

XX Haemostatic; antihaemic; immunosuppressive; platelet;

KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;

KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KW autoimmune haemolytic anaemia; Hughes' syndrome;

XX lupoid thrombocytopenia.

OS Homo sapiens.

XX WO2003031589-A2.

XX 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

PA (AMGE-) AMGEN INC.

PI Min H, Stiney KC, Hartley C;

DR WPI; 2003-403101/38.

DR N-PSDB; ADN59707.

XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

PT which stimulate the production of platelets and/or the production of

DR platelet precursors, useful for treating thrombocytopenia.

XX Disclosure; SEQ ID NO 57; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
 CC platelets and/or the production of platelet precursors, is new. Further  
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
 CC and a pharmaceutical composition comprising (II) and a carrier. The  
 CC pharmaceutical composition of the invention is useful for treating  
 CC thrombocytopenia in an animal, and for increasing megakaryocytes or





CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide.

XX  
SQ Sequence 15 AA;

Query Match 67.3%; Score 66; DB 8; Length 15;  
Best Local Similarity 73.3%; Pred. No. 0.0013;  
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 IEQPTLRWLTSTRTP 16  
| | | | | : | |  
| | | | | : | |  
Db 1 IEQPTLRWLTAAARAP 15

#### RESULT 22

ABP51687  
ID ABP51687 standard; peptide; 18 AA.

XX  
AC ABP51687;

XX  
DT 01-OCT-2002 (first entry)

XX  
DE TPO mimetic peptide SEQ ID NO:37.

XX  
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
XX complementarily determining region; immunoglobulin; antianemic;  
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX  
OS Homo sapiens.

OS Synthetic.

XX  
PN WO200246238-A2.

XX  
PD 13-JUN-2002.

XX  
PF 05-DEC-2001; 2001WO-US047656.

XX  
PR 05-DEC-2000; 2000US-0251448P.

XX  
PR 04-MAY-2001; 2001US-0288889P.

XX  
PR 29-MAY-2001; 2001US-0294068P.

XX  
PA (ALEX-) ALEXION PHARM INC.

XX  
PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX  
DR WPI; 2002-566610/60.

XX  
DR N-PSDB; ABQ73365.

XX  
PT A novel immunogen molecule comprising a region in which amino acid  
XX residues corresponding to at least a portion of the complementary  
XX determining region are replaced or fused with an erythropoietin or  
XX thrombopoietin mimetic.

XX  
PS Claim 20; Fig 5; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment  
CC (I) comprising a region where amino acid residues corresponding to at  
CC least a portion of the complementary determining region (CDR) are  
CC replaced or fused with biologically active peptides e.g. a peptide  
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
CC that is flanked with proline at its carboxy terminus. (I) has  
CC antianemic, haemostatic and nephrotropic activities, and can be used as  
CC a stimulator of proliferation, differentiation and maturation of  
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
CC for stimulating proliferation, differentiation or growth of  
CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
CC promegakaryocytes or megakaryocytes, which results in increased platelet  
CC production. (I) with a region where amino acid residues corresponding to  
CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
CC production of red blood cells, where (I) is contacted with haematopoietic  
CC stem cells or their progenitors. (I) is useful for diagnostics or

CC therapeutics, in cell isolation strategies, and for treating patients  
CC suffering from deficiency in cell populations caused by disease,  
CC disorders or treatments related to the suppression of haematopoiesis.  
CC ABQ73288 to ABQ73377 and ABP51689 to ABP51696 represent sequences used in  
CC the exemplification of the present invention

XX  
SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
Best Local Similarity 73.3%; Pred. No. 0.0016;  
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 IEQPTLRWLTSTRTP 16  
| | | | | : | |  
| | | | | : | |  
Db 3 IEQPTLRWLTAAARAP 17

#### RESULT 23

ABP51689  
ID ABP51689 standard; peptide; 18 AA.

XX  
AC ABP51689;

XX  
DT 01-OCT-2002 (first entry)

XX  
DE TPO mimetic peptide SEQ ID NO:41.

XX  
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
XX complementarily determining region; immunoglobulin; antianemic;  
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX  
OS Homo sapiens.

OS Synthetic.

XX  
PN WO200246238-A2.

XX  
PD 13-JUN-2002.

XX  
PF 05-DEC-2001; 2001WO-US047656.

XX  
PR 05-DEC-2000; 2000US-0251448P.

XX  
PR 04-MAY-2001; 2001US-0288889P.

XX  
PR 29-MAY-2001; 2001US-0294068P.

XX  
PA (ALEX-) ALEXION PHARM INC.

XX  
PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX  
DR WPI; 2002-566610/60.

XX  
DR N-PSDB; ABQ73367.

XX  
PT A novel immunogen molecule comprising a region in which amino acid  
XX residues corresponding to at least a portion of the complementary  
XX determining region are replaced or fused with an erythropoietin or  
XX thrombopoietin mimetic.

XX  
PS Claim 20; Fig 5; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment  
CC (I) comprising a region where amino acid residues corresponding to at  
CC least a portion of the complementary determining region (CDR) are  
CC replaced or fused with biologically active peptides e.g. a peptide  
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
CC that is flanked with proline at its carboxy terminus. (I) has  
CC antianemic, haemostatic and nephrotropic activities, and can be used as  
CC a stimulator of proliferation, differentiation and maturation of  
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
CC for stimulating proliferation, differentiation or growth of  
CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
CC promegakaryocytes or megakaryocytes, which results in increased platelet  
CC production. (I) with a region where amino acid residues corresponding to  
CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51689 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 IEGPTLRWLTGRTP 16  
 |||||:|:|:  
 Db 3 IEGPTLRQWLARAP 17

# RESULT 24

ABP51688 ID ABP51688 standard; peptide; 18 AA.

XX AC ABP51688;

XX DT 01-OCT-2002 (first entry)

XX DE TPO mimetic peptide SEQ ID NO:39.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.  
 OS Synthetic.

XX PN WO200246238-A2.

XX PD 13-JUN-2002.

XX PF 05-DEC-2001; 2001WO-US047656.

XX PR 05-DEC-2000; 2000US-0251448P.

XX PR 04-MAY-2001; 2001US-0288889P.

XX PR 29-MAY-2001; 2001US-0294068P.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX DR WPI; 2002-566610/60.

XX DR N-PSDB; ABQ73366.

XX PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

XX PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51689 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 IEGPTLRWLTGRTP 16  
 |||||:|:|:  
 Db 3 IEGPTLRQWLARAP 17

# RESULT 25

ABP51686 ID ABP51686 standard; peptide; 18 AA.

XX AC ABP51686;

XX DT 01-OCT-2002 (first entry)

XX DE TPO mimetic peptide SEQ ID NO:35.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.  
 OS Synthetic.

XX PN WO200246238-A2.

XX PD 13-JUN-2002.

XX PF 05-DEC-2001; 2001WO-US047656.

XX PR 05-DEC-2000; 2000US-0251448P.

XX PR 04-MAY-2001; 2001US-0288889P.

XX PR 29-MAY-2001; 2001US-0294068P.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX DR WPI; 2002-566610/60.

XX DR N-PSDB; ABQ73364.

XX PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

XX PS Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is useful for diagnostics or  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease.  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP5169 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGLPTLRWLTSTRTP 16  
 |||||:|:|:  
 Db 3 IEGLPTLRQWLAAAP 17

## RESULT 26

ABP51684  
 ID ABP51684 standard; peptide; 18 AA.

AC ABP51684;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:31.

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

DR WPI; 2002-566610/60.

DR N-PSDB; ABQ73362.

PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

PS Claim 20; Fig 5; 113pp; English.

XX  
 XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is useful for diagnostics or  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease.  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP5169 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGLPTLRWLTSTRTP 16  
 |||||:|:|:  
 Db 3 IEGLPTLRQWLAAAP 17

## RESULT 27

ABP51690  
 ID ABP51690 standard; peptide; 18 AA.

AC ABP51690;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:43.

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

DR WPI; 2002-566610/60.

DR N-PSDB; ABQ73368.

PT A novel immunogen molecule comprising a region in which amino acid  
 PT residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.

PS Claim 20; Fig 5; 113pp; English.

XX  
 XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells; and a stimulator of haematopoiesis. (1) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (1) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (1) is contacted with haematopoietic  
 CC stem cells or their progenitors. (1) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP5169 to ABP5166 represent sequences used in  
 CC the exemplification of the present invention

CC SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016; 2; Mismatches 2; Indels 0; Gaps 0;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTPT 16  
 |||||:|:|  
 Db 3 IEGPTLRWLTSTPT 17

RESULT 28

ABP51675  
 ID ABP51675 standard; peptide; 18 AA.

AC ABP51675;

DT 01-OCT-2002 (first entry)

DE TPO mimetic antibody related peptide graft SEQ ID NO:66.

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

KM complementarity determining region; immunoglobulin; antianaemic;

KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

OS Synthetic.

PN WO200246238-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US047656.

PR 05-DEC-2000; 2000US-0251448P.

PR 04-MAY-2001; 2001US-0288889P.

PR 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

PA Bowdish KS, Barbas-Frederickson S, Renshaw M;

PI WPI; 2002-566610/60.

XX A novel immunogen molecule comprising a region in which amino acid

PT residues corresponding to at least a portion of the complementary

PT determining region are replaced or fused with an erythropoietin or

PT thrombopoietin mimetic.

XX Example 4; Page 55; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment

CC (1) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (1) has

CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells; and a stimulator of haematopoiesis. (1) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (1) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (1) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (1) is contacted with haematopoietic  
 CC stem cells or their progenitors. (1) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP5169 to ABP5166 represent sequences used in  
 CC the exemplification of the present invention

CC SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016; 2; Mismatches 2; Indels 0; Gaps 0;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTPT 16  
 |||||:|:|  
 Db 3 IEGPTLRWLTSTPT 17

RESULT 29

ADQ16611  
 ID ADQ16611 standard; peptide; 18 AA.

AC ADQ16611;

DT 09-SEP-2004 (first entry)

DE TPO mimetic peptide with random flanking residues SEQ ID NO:31.

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;

KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;

KW immunotherapy; thrombocytopenia.

OS Unidentified.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

PR (ALEX-) ALEXION PHARM INC.

PA Bowdish KS, Frederickson S, Renshaw M;

PI WPI; 2004-460973/43.

XX N-PSDB; ADQ16612.

PT New immunoglobulin molecule comprising a region, where two

PT complementarity determining regions (CDRs) are replaced with EPO mimetic

PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 1; SEQ ID NO 31; 107pp; English.

CC The invention relates to a novel immunoglobulin molecule or its fragment

CC comprising a region where amino acid residues corresponding to at least a

CC portion of a two complementarity determining regions (CDRs) are replaced

CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and

CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the

CC invention has immunosuppressive activity, and may have a use in

CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or

CC treating thrombocytopenia as a result of chemotherapy, bone marrow

CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

XX  
 SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16  
 |||||:|:|  
 Db 3 IEGPTLRQWLARAP 17

RESULT 30

ADQ16619  
 ID ADQ16619 standard; peptide; 18 AA.

AC ADQ16619;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide with random flanking residues SEQ ID NO:39.

DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.

XX Unidentified.

OS WO2004050017-A2.

PN 17-JUN-2004.

PD 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

PR (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;

PI WPI; 2004-460973/43.

XX N-PSDB; ADQ16620.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 1; SEQ ID NO 39; 107pp; English.

PS The invention relates to a novel immunoglobulin molecule or its fragment  
 XX comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplant, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

XX Sequence 18 AA;

Query Match 67.3%; Score 66; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16  
 |||||:|:|  
 Db 3 IEGPTLRQWLARAP 17

RESULT 31

ADQ16621  
 ID ADQ16621 standard; peptide; 18 AA.

AC ADQ16621;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide with random flanking residues SEQ ID NO:41.

DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.

XX Unidentified.

OS WO2004050017-A2.

PN 17-JUN-2004.

PD 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

PR (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;

PI WPI; 2004-460973/43.

XX N-PSDB; ADQ16622.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Example 1; SEQ ID NO 41; 107pp; English.

PS The invention relates to a novel immunoglobulin molecule or its fragment  
 XX comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplant, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

XX Sequence 18 AA;

Query Match 67.3%; Score 66; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16  
 |||||:|:|  
 Db 3 IEGPTLRQWLARAP 17

RESULT 32

ADQ16646  
 ID ADQ16646 standard; peptide; 18 AA.

AC ADQ16646;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide SEQ ID NO:65.

DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW immunotherapy; thrombocytopenia.

KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.  
 XX Unidentified.  
 XX MO2004050017-A2.  
 XX 17-JUN-2004.  
 XX 17-NOV-2003; 2003WO-US036894.  
 XX 02-DEC-2002; 2002US-00307724.  
 XX (ALEX-) ALEXION PHARM INC.  
 XX Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.  
 XX N-PSDB; ADQ16645.  
 XX New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX Example 4; SEQ ID NO 66; 107pp; English.  
 XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide of the invention.  
 XX Sequence 18 AA;  
 XX SQ

Query Match 67.3%; Score 66; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16  
 |||||:|:|:  
 DB 3 IEGPTLRQWTLARAP 17

RESULT 33  
 ADQ16615  
 ID ADQ16615 standard; peptide; 18 AA.  
 XX ADQ16615;  
 AC ADQ16615;  
 XX 09-SEP-2004 (first entry)  
 DT TPO mimetic peptide with random flanking residues SEQ ID NO:35.  
 XX DE  
 XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.  
 XX Unidentified.  
 XX OS  
 XX WO2004050017-A2.  
 XX 17-JUN-2004.  
 XX 17-NOV-2003; 2003WO-US036894.  
 XX 02-DEC-2002; 2002US-00307724.  
 XX (ALEX-) ALEXION PHARM INC.  
 XX Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.  
 XX N-PSDB; ADQ16615.

XX Bowdish KS, Frederickson S, Renshaw M;  
 PI WPI; 2004-460973/43.  
 XX N-PSDB; ADQ16615.  
 XX New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX Example 1; SEQ ID NO 35; 107pp; English.  
 XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.  
 XX Sequence 18 AA;  
 XX SQ

Query Match 67.3%; Score 66; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.0016;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16  
 |||||:|:|:  
 DB 3 IEGPTLRQWTLARAP 17

RESULT 34  
 ADQ16617  
 ID ADQ16617 standard; peptide; 18 AA.  
 XX ADQ16617;  
 AC ADQ16617;  
 XX 09-SEP-2004 (first entry)  
 DT TPO mimetic peptide with random flanking residues SEQ ID NO:37.  
 XX DE  
 XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.  
 XX Unidentified.  
 XX OS  
 XX WO2004050017-A2.  
 XX 17-JUN-2004.  
 XX 17-NOV-2003; 2003WO-US036894.  
 XX 02-DEC-2002; 2002US-00307724.  
 XX (ALEX-) ALEXION PHARM INC.  
 XX Bowdish KS, Frederickson S, Renshaw M;  
 XX WPI; 2004-460973/43.  
 XX N-PSDB; ADQ16618.  
 XX New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX Example 1; SEQ ID NO 37; 107pp; English.  
 XX The invention relates to a novel immunoglobulin molecule or its fragment



CC comprising a region where amino acid residues corresponding to at least a  
CC portion of a two complementarity determining regions (CDRs) are replaced  
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
CC invention has immunosuppressive activity, and may have a use in  
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide with flanking  
CC residues.  
CC

SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 8; Length 18;  
Best Local Similarity 73.3%; Pred. No. 0.0016;  
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTP 16  
|||||:|:|  
DB 3 IEGPLROWLAARAP 17

RESULT 35

ADQ16623  
ID ADQ16623 standard; peptide; 18 AA.

AC ADQ16623;

DT 09-SEP-2004 (first entry)

DE TPO mimetic peptide with random flanking residues SEQ ID NO:43.

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
KW immunotherapy; thrombocytopenia.

OS Unidentified.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;

DR WPI; 2004-460973/43.

DR N-PSDB; ADQ16624.

PT New immunoglobulin molecule comprising a region, where two  
PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
PT or a TPO mimetic, useful for treating thrombocytopenia.  
XX

PS Example 1; SEQ ID NO 43; 107bp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
CC comprising a region where amino acid residues corresponding to at least a  
CC portion of a two complementarity determining regions (CDRs) are replaced  
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
CC invention has immunosuppressive activity, and may have a use in  
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents a TPO mimetic peptide with flanking  
CC residues.  
CC

SQ Sequence 18 AA;

Query Match 67.3%; Score 66; DB 8; Length 18;  
Best Local Similarity 73.3%; Pred. No. 0.0016;  
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTP 16  
|||||:|:|  
DB 3 IEGPLROWLAARAP 17

RESULT 36

ADQ16708  
ID ADQ16708 standard; protein; 22 AA.

AC ADQ16708;

DT 09-SEP-2004 (first entry)

DE Immunoglobulin heavy chain CDR2 with TPO clone HR2-20 SEQ ID NO:128.

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
KW immunotherapy; thrombocytopenia.

OS Synthetic.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;

DR WPI; 2004-460973/43.

PT New immunoglobulin molecule comprising a region, where two  
PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
PT or a TPO mimetic, useful for treating thrombocytopenia.  
XX

PS Example 9; SEQ ID NO 128; 107bp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
CC comprising a region where amino acid residues corresponding to at least a  
CC portion of a two complementarity determining regions (CDRs) are replaced  
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
CC invention has immunosuppressive activity, and may have a use in  
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
CC The present sequence represents an immunoglobulin heavy chain CDR2 with  
CC TPO peptide inserted.  
XX

SQ Sequence 22 AA;

Query Match 67.3%; Score 66; DB 8; Length 22;  
Best Local Similarity 68.8%; Pred. No. 0.0021;  
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTPH 17  
|||||:|:|  
DB 6 IEGPLROWLAARAH 21

RESULT 37

ADQ16710  
ID ADQ16710 standard; protein; 22 AA.

AC ADQ16710;



```
XX 09-SEP-2004 (first entry)
DT Immunoglobulin heavy chain CDR2 with TPO clone HR2-28 SEQ ID NO:130.
XX
XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;
DE erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
XX immunotherapy; thrombocytopenia.
XX
XX Synthetic.
XX
XX WO2004050017-A2.
XX
XX 17-JUN-2004.
XX
XX 17-NOV-2003; 2003WO-US036894.
XX
XX 02-DEC-2002; 2002US-00307724.
XX
XX (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Frederickson S, Renshaw M;
XX
XX WPI; 2004-460973/43.
XX
XX New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
XX Example 9; SEQ ID NO 130; 107pp; English.
XX
XX The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents an immunoglobulin heavy chain CDR2 with
CC TPO peptide inserted.
XX
XX Sequence 22 AA;
SQ
Query Match 67.3%; Score 66; DB 8; Length 22;
Best Local Similarity 73.3%; Pred. No. 0.0021;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 IEGPTLRWLTSTRTP 16
| | | | | | | | | | | |
| | | | | | | | | | | |
Db 6 IEGPTLRQWLARAP 20
| | | | | | | | | | | |
| | | | | | | | | | | |
RESULT 38
ADQ16705
ID ADQ16705 standard; protein; 128 AA.
XX
XX ADQ16705;
XX
XX 09-SEP-2004 (first entry)
XX
XX Modified immunoglobulin clone 116 HC variable region SEQ ID NO:125.
XX
XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
XX immunotherapy; thrombocytopenia.
XX
XX Synthetic.
XX
XX WO2004050017-A2.
XX
XX 17-JUN-2004.
PD
```

```
XX 17-NOV-2003; 2003WO-US036894.
XX
XX 02-DEC-2002; 2002US-00307724.
XX
XX (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Frederickson S, Renshaw M;
XX
XX WPI; 2004-460973/43.
XX
XX New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
XX Claim 9; SEQ ID NO 125; 107pp; English.
XX
XX The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents immunoglobulin clone 116 heavy chain
CC variable region.
XX
XX Sequence 128 AA;
SQ
Query Match 67.3%; Score 66; DB 8; Length 128;
Best Local Similarity 73.3%; Pred. No. 0.015;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 IEGPTLRWLTSTRTP 16
| | | | | | | | | | | |
| | | | | | | | | | | |
Db 102 IEGPTLRQWLARAP 116
| | | | | | | | | | | |
| | | | | | | | | | | |
RESULT 39
ADQ16704
ID ADQ16704 standard; protein; 225 AA.
XX
XX ADQ16704;
XX
XX 09-SEP-2004 (first entry)
XX
XX Modified immunoglobulin clone 116 heavy chain SEQ ID NO:124.
XX
XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
XX immunotherapy; thrombocytopenia.
XX
XX Synthetic.
XX
XX WO2004050017-A2.
XX
XX 17-JUN-2004.
XX
XX 17-NOV-2003; 2003WO-US036894.
XX
XX 02-DEC-2002; 2002US-00307724.
XX
XX (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Frederickson S, Renshaw M;
XX
XX WPI; 2004-460973/43.
XX
XX New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
```

XX Example 8; SEQ ID NO 124; 107pp; English.

PS The invention relates to a novel immunoglobulin molecule or its fragment

XX comprising a region where amino acid residues corresponding to at least a

CC portion of a two complementarity determining regions (CDRs) are replaced

CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and

CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the

CC invention has immunosuppressive activity, and may have a use in

CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or

CC treating thrombocytopenia as a result of chemotherapy, bone marrow

CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.

CC The present sequence represents immunoglobulin clone 116 heavy chain.

XX Sequence 225 AA;

SO

Query Match 67.3%; Score 66; DB 8; Length 225;

Best Local Similarity 73.3%; Pred. No. 0.028;

Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16

DB 102 IEGPTLRQWLAKAP 116

RESULT 40

ABP51695

ID ABP51695 standard; protein; 472 AA.

XX

AC ABP51695;

XX

DT 01-OCT-2002 (first entry)

XX

DE 5G1.1-TPO heavy chain amino acid sequence SEQ ID NO:67.

XX

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

KM complementarity determining region; immunoglobulin; antianaemic;

XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

OS Synthetic.

XX

PN WO200246238-A2.

XX

PD 13-JUN-2002.

XX

PF 05-DEC-2001; 2001WO-US047656.

XX

PR 05-DEC-2000; 2000US-0251448P.

XX

PR 04-MAY-2001; 2001US-0288889P.

XX

PR 29-MAY-2001; 2001US-0294068P.

XX

PA (ALEX-) ALEXION PHARM INC.

XX

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX

PI WPI: 2002-566610/60.

XX

DR N-PSDB; ABO73374.

XX

PT A novel immunogen molecule comprising a region in which amino acid

PT residues corresponding to at least a portion of the complementary

PT determining region are replaced or fused with an erythropoietin or

PT thrombopoietin mimetic.

XX

XX Example 4; Fig 13A; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment

CC (I) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (I) has

CC antianaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of

CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or

CC therapeutics, in cell isolation strategies, and for treating patients

CC suffering from deficiency in cell populations caused by disease,

CC disorders or treatments related to the suppression of haematopoiesis.

CC ABO73288 to ABO73377 and ABP51695 to ABP51696 represent sequences used in

CC the exemplification of the present invention

XX

SO Sequence 472 AA;

Query Match 67.3%; Score 66; DB 5; Length 472;

Best Local Similarity 73.3%; Pred. No. 0.063;

Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTP 16

DB 120 IEGPTLRQWLAKAP 134

RESULT 41

ADQ16647

ID ADQ16647 standard; protein; 472 AA.

XX

AC ADQ16647;

XX

DT 09-SEP-2004 (first entry)

XX

DE Immunoglobulin antibody 5G1.1-TPO heavy chain SEQ ID NO:67.

XX

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;

KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;

XX immunotherapy; thrombocytopenia.

OS Synthetic.

OS

PN WO2004050017-A2.

XX

PD 17-JUN-2004.

XX

PF 17-NOV-2003; 2003WO-US036894.

XX

PR 02-DEC-2002; 2002US-00307724.

XX

PR (ALEX-) ALEXION PHARM INC.

XX

PI Bowdish KS, Frederickson S, Renshaw M;

XX

PI WPI: 2004-460973/43.

XX

DR N-PSDB; ADQ16648.

XX

PT New immunoglobulin molecule comprising a region, where two

PT complementarity determining regions (CDRs) are replaced with EPO mimetic

PT or a TPO mimetic, useful for treating thrombocytopenia.

XX

XX Example 4; SEQ ID NO 67; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment

CC comprising a region where amino acid residues corresponding to at least a

CC portion of a two complementarity determining regions (CDRs) are replaced

CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and

CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the

CC invention has immunosuppressive activity, and may have a use in

CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or

CC treating thrombocytopenia as a result of chemotherapy, bone marrow

CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.

CC The present sequence represents an immunoglobulin antibody heavy chain of  
 CC the invention.  
 XX  
 SQ Sequence 472 AA;  
 Query Match 67.3%; Score 66; DB 8; Length 472;  
 Best Local Similarity 73.3%; Pred. No. 0.063;  
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 IEGPTLEWLTSTRT 16  
 DB 120 IEGPTLRQWLAARAP 134  
 RESULT 42  
 AAB16969  
 ID AAB16969 standard; peptide; 14 AA.  
 AC AAB16969;  
 DT 31-OCT-2000 (first entry)  
 DE TPO-mimetic peptide sequence SEQ ID NO:25.  
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytotoxic; antineoplastic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.  
 OS Synthetic.  
 PN MO200024782-A2.  
 PD 04-MAY-2000.  
 PF 25-OCT-1999; 99MO-US025044.  
 PR 23-OCT-1998; 98US-0105371P.  
 PR 22-OCT-1999; 99US-00428082.  
 PA (AMGE-) AMGEN INC.  
 PI Feige U, Liu C, Cheatham J, Boone TC;  
 DR WPI; 2000-350702/30.  
 PT Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 PS Claim 19; Page 203; 608pp; English.  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-E-F1-(X2)-b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-C-P1, -(L1)-C-P1-(L2)-d-P2, -(L1)-C-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytotoxic, antineoplastic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 XX

SQ Sequence 14 AA;  
 Query Match 63.3%; Score 62; DB 3; Length 14;  
 Best Local Similarity 84.6%; Pred. No. 0.0055;  
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 IEGPTLEWLTSTRT 14  
 DB 1 IEGPTLRQWLAAR 13  
 RESULT 43  
 ABB72855  
 ID ABB72855 standard; peptide; 14 AA.  
 AC ABB72855;  
 DT 05-APR-2002 (first entry)  
 DE TPO mimetic peptide SEQ ID NO:25.  
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IGG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antineoplastic; antitumour; immunosuppressive;  
 KW cytotoxic; antineoplastic; antitumour; antidiabetic; ophthalmological;  
 KW antineoplastic; anorectic; antinfertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO2001B3525-A2.  
 PD 08-NOV-2001.  
 PF 02-MAY-2001; 2001WO-US014310.  
 PR 03-MAY-2000; 2000US-00563286.  
 PA (AMGE-) AMGEN INC.  
 PI Feige U, Liu C, Cheatham JC, Boone TC, Gudae JM;  
 DR WPI; 2002-130313/17.  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 43; 176pp; English.  
 CC The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytotoxic, antineoplastic, antitumour, antidiabetic, ophthalmological and  
 CC antineoplastic, anorectic, antinfertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopenia, aplastic anemia, metastatic  
CC tumor which result in thrombocytopenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777  
CC represent amino acid and nucleic acid sequences used in the  
CC exemplification of the present invention

XX Sequence 14 AA;

SO Query Match 63.3%; Score 62; DB 5; Length 14;  
Best Local Similarity 84.6%; Pred. No. 0.0055;  
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 2 IEPTLRRLTISR 14  
| | | | | | | | | | | | : |  
Db 1 IEPTLRRLTISR 13

RESULT 44

ID ADJ73005 standard; peptide; 14 AA.

AC ADJ73005;

DT 06-MAY-2004 (first entry)

DE TPO mimetic peptide sequence SegID 459.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
KM immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
KM TPO.

XX Synthetic.

OS WO2003084477-A2.

PN 16-OCT-2003.

PD 24-MAR-2003; 2003WO-US009139.

PF 29-MAR-2002; 2002US-0368791P.

PR (CENZ) CENTOCOR INC.

PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain  
PT variable region comprising human framework or ligand binding region,  
PT useful for preparing a composition for treating e.g., immune,  
PT cardiovascular or neurologic disease.

PS Disclosure; SEQ ID NO 459; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific  
CC portions or variants thereof. Specifically, it refers to an antibody  
CC fragment where a protein has been inserted into, or replaces a portion  
CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
CC least one portion of a heavy chain or light chain variable region, which  
CC itself comprises at least one human framework region and at least one  
CC ligand binding region (LBR). The present invention describes human  
CC mimetibodies, including modified immunoglobulins and cleavage products  
CC that can be useful in gene therapy and the generation of transgenic  
CC plants and animals. Furthermore, the CDR mimetibody is useful for  
CC preparing compositions for modulating, treating or reducing the symptoms  
CC of immune, cardiovascular, infectious, malignant and/or neurologic  
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
CC peptide sequence is a TPO mimetic peptide sequence used to make a  
CC mimetibody of the invention.

SO Sequence 14 AA;

Query Match 63.3%; Score 62; DB 7; Length 14;  
Best Local Similarity 84.6%; Pred. No. 0.0055;  
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 IEPTLRRLTISR 14  
| | | | | | | | | | | | : |  
Db 1 IEPTLRRLTISR 13

RESULT 45

ID ADJ52640 standard; peptide; 14 AA.

AC ADJ52640;

DT 06-MAY-2004 (first entry)

DE CHI deleted mimetibody-related peptide SegID459.

XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
KM hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
KM fungicide; gene therapy; immune disorder; cardiovascular disease;  
KM arrhythmia; hypertension; heart failure; neurodegenerative;  
KM multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
KM cancerous condition; infectious disease; bacterial infection;  
KM viral infection; fungal infection.

OS Unidentified.

OS Synthetic.

PN WO2004002417-A2.

PD 08-JAN-2004.

PF 27-JUN-2003; 2003WO-US020347.

PR 28-JUN-2002; 2002US-0392431P.

PA (CENZ) CENTOCOR INC.

PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neseppor TC,  
XX Kutooski KA;

XX WPI; 2004-082870/08.

XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
PT diseases.

PS Claim 2; SEQ ID NO 459; 129pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an immunosuppressive,  
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
CC is useful for diagnosing or treating a disease condition in a cell,  
CC tissue, organ or animal, specifically for modulating, treating,  
CC alleviating, preventing the incidence or reducing the symptoms of an  
CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
CC conditions, or infectious diseases (for example bacterial, viral or  
CC fungal infection). The present sequence is that of a peptide which may be  
CC used during the creation of a mimetibody of the invention.

XX Sequence 14 AA;

SO Query Match 63.3%; Score 62; DB 8; Length 14;

	Best Local Similarity	84.6%	Pred. NO. 0.0055			
Matches	11	Conservative	1	Mismatches	1	Indels
						Gaps
Oy	2	IEGPTLRWMLTSR	14			
			:			
			:			
Db	1	IEGPTLRWMLAAR	13			

Search completed: September 1, 2005, 16:12:12  
Job time : 82.7482 BECS

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:57:33 ; Search time 13.7266 Seconds  
 (without alignments)  
 126.171 Million cell updates/sec

Title: US-10-083-768-10

Perfect score: 98  
 Sequence: 1 SIEGPTLRKWLRTSRPHS 18

Scoring table: BIOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 100 summaries

Database : PIR 79:\*  
 1: p1r1:\*  
 2: p1r2:\*  
 3: p1r3:\*  
 4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	52.0	753	1	D72660
2	48.5	49.5	410	2	DEPSXA
3	48.5	49.5	410	2	C83365
4	48	49.0	150	2	C75456
5	48	49.0	664	2	G89894
6	47	48.0	154	2	F64026
7	47	48.0	430	2	B69659
8	46.5	47.4	330	2	C69593
9	46.5	47.4	527	2	B64633
10	46	46.9	664	2	H83962
11	46	46.9	1028	2	A59253
12	45.5	46.4	331	2	AD1246
13	45.5	46.4	331	2	AH1608
14	45	45.9	132	2	S15618
15	45	45.9	473	2	B84853
16	45	45.9	886	2	A32758
17	45	45.9	1036	2	S22383
18	44.5	45.9	2354	2	T13288
19	44	44.9	200	2	T23485
20	44	44.9	207	2	T37464
21	44	44.9	310	2	C90277
22	44	44.9	400	2	C87021
23	43.5	44.4	327	2	B82277
24	43	43.9	347	2	T06371
25	43	43.9	399	2	B70936
26	43	43.9	491	2	C98275
27	43	43.9	491	2	AC3009
28	43	43.9	514	2	T44502
29	43	43.9	648	1	H69878

30	43	43.9	1094	2	F70697
31	43	43.9	2896	2	T30939
32	42.5	43.4	371	1	DEBSPA
33	42.5	43.4	436	2	JC4742
34	42.5	43.4	586	2	B84271
35	42	42.9	263	2	AH1884
36	42	42.9	353	2	G87394
37	42	42.9	306	2	D70601
38	42	42.9	310	2	JL0119
39	42	42.9	317	2	JL0118
40	42	42.9	323	2	S06946
41	42	42.9	430	2	B82096
42	42	42.9	448	2	B45438
43	42	42.9	564	2	T37934
44	42	42.9	721	2	A39707
45	42	42.9	807	2	H75634
46	42	42.9	970	2	A72028
47	42	42.9	970	2	G86595
48	42	42.9	1028	2	S41749
49	42	42.9	1028	2	S37146
50	41.5	42.3	125	1	A46315
51	41.5	42.3	330	2	C83995
52	41.5	42.3	4006	2	T09070
53	41	41.8	109	2	S69853
54	41	41.8	285	2	AC1537
55	41	41.8	306	2	T45453
56	41	41.8	336	2	S41643
57	41	41.8	346	2	B84377
58	41	41.8	346	2	B90073
59	41	41.8	389	2	B69096
60	41	41.8	410	2	G90362
61	41	41.8	413	2	A10598
62	41	41.8	482	2	D75346
63	41	41.8	521	2	T01923
64	41	41.8	754	2	D88734
65	41	41.8	897	2	B69202
66	41	41.8	1019	2	T11560
67	41	41.8	1040	2	T08190
68	41	41.8	1123	2	T51517
69	41	41.8	1172	2	AD2310
70	41	41.8	1299	2	A42090
71	40.5	41.3	255	2	A45881
72	40.5	41.3	371	2	AD1206
73	40.5	41.3	371	2	AC1563
74	40.5	41.3	446	1	IOB80C
75	40.5	41.3	1420	2	A32869
76	40.5	41.3	2476	2	T34022
77	40	40.8	195	2	B56688
78	40	40.8	195	2	A85481
79	40	40.8	195	2	A90630
80	40	40.8	195	2	AE0057
81	40	40.8	220	2	AC0318
82	40	40.8	250	2	T14548
83	40	40.8	251	2	F95295
84	40	40.8	279	2	G83041
85	40	40.8	289	2	J00059
86	40	40.8	312	2	F86876
87	40	40.8	327	2	B71900
88	40	40.8	331	2	B48445
89	40	40.8	337	1	DEJUGC
90	40	40.8	338	2	T47218
91	40	40.8	378	2	D83381
92	40	40.8	413	2	H75357
93	40	40.8	466	2	B86411
94	40	40.8	492	2	C84142
95	40	40.8	524	2	E71881
96	40	40.8	561	2	C75543
97	40	40.8	694	2	A96571
98	40	40.8	719	2	B95325
99	40	40.8	722	2	T37970
100	40	40.8	749	2	H82691

probable arabinosy  
 hemocyanin G-type  
 pyruvate dehydroge  
 transposase - Cory  
 glutamyl-LRNA synt  
 hypothetical prote  
 hypothetical prote  
 UTP-glucose-1-phos  
 Fc gamma (IgG) rec  
 Fc gamma (IgG) rec  
 Fc gamma (IgG) rec  
 myosin I beta, PMI  
 conserved hypotet  
 erythrocyte membra  
 myosin-1C - mouse  
 preprotein translo  
 protein translocas  
 myosin heavy chain  
 myosin I heavy cha  
 E4 protein - human  
 branched-chain alp  
 probable tenascin  
 hypothetical prote  
 hypothetical prote  
 UTP-glucose-1-phos  
 svrm protein - Rh1  
 protein export [im  
 hypothetical prote  
 cortinoid/iron-sul  
 hypothetical prote  
 probable phosphol  
 glucamyl-LRNA(Gln)  
 hypothetical prote  
 protein F32P10.1 l  
 valine-tRNA ligase  
 pol polyprotein -  
 hypothetical prote  
 telomerase reverse  
 hypothetical prote  
 two-component hybr  
 MHC class II histo  
 pyruvate dehydroge  
 pyruvate dehydroge  
 replication initia  
 apolipoprotein(a)  
 zonahesin - pig  
 molybdopterin bios  
 molybdopterin bio  
 molybdopterin bio  
 molybdopterin bios  
 probable nicotinat  
 beta-fructofuranos  
 gluconate 5-dehydr  
 probable N-hydroxy  
 hypothetical 31.6K  
 hypothetical prote  
 glyceraldehyde-3-P  
 glyceraldehyde-3-P  
 hypothetical prote  
 tRNA (5-methylamin  
 protein P3M18.4 [i  
 hypothetical prote  
 hypothetical prote  
 6-aminohexanoate-c  
 hypothetical prote  
 conserved hypotet  
 probable G2-specif  
 topoisomerase IV 8

## ALIGNMENTS

```

RESULT 1
D72660
probable aldehyde oxidoreductase APE0708 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: D72660
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72660
A:Molecule type: DNA
A:Residues: 1-753 <KAW>
A:Cross-references: UNIPROT:Q9Y6E2; DDBJ:AB000060; NID:G5104188; PIDN:BAA79684.1; PID:G5
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0708
C:Superfamily: carbon monoxide dehydrogenase molybdoprotein

Query Match          52.0%; Score 51; DB 1; Length 753;
Best Local Similarity 47.1%; Pred. No. 4.7;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 SIEGPTLRRLTSTRTPH 17
      ||:|||||:|:|:|
Db      192 SYDGSRLTWVSTQTPH 208

RESULT 2
DEPSXA
3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) alpha chain - Pseudomonas
N:Alternate names: 2-oxoisovalerate dehydrogenase (lipoamide) E1-alpha chain; branched-C
C:Species: Pseudomonas putida
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: S01317; B36133; S63475
R:Burns, G.; Brown, T.; Halter, K.; Idriss, J.M.; Sokatch, J.R.
Eur. J. Biochem. 176, 311-317, 1988
A:Title: Similarity of the E1 subunits of branched-chain-oxoacid dehydrogenase from Pseu
A:Reference number: S01317; MUID:88329084; PMID:3416875
A:Accession: S01317
A:Molecule type: DNA
A:Residues: 1-410 <BUR>
A:Cross-references: UNIPROT:P09060; EMBL:X13004
R:Madhusudan, K.T.; Huang, G.; Burns, G.; Sokatch, J.R.
J. Bacteriol. 172, 5655-5663, 1990
A:Title: Transcriptional analysis of the promoter region of the Pseudomonas putida bran
A:Reference number: A36133; MUID:9100895; PMID:2211503
A:Accession: B36133
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-17 <MAD>
A:Cross-references: GB:M33715
R:Heester, K.; Luo, J.; Burns, G.; Braswell, E.H.; Sokatch, J.R.
Eur. J. Biochem. 233, 828-836, 1995
A:Title: Purification of active E1-alpha(2)-beta(2) of Pseudomonas putida branched-chain
A:Reference number: S63475; MUID:96085147; PMID:8521848
A:Accession: S63475
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <HES>
C:Genetics:
A:Gene: bkdA1
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin
C:Keyword: lipoamide; oxidoreductase; phosphoprotein; thiamin pyrophosphate
P:2-410/Product: 3-methyl-2-oxobutanoate dehydrogenase (lipoamide) alpha chain #status F
F:202-251/Domain: thiamin pyrophosphate-binding domain homology <TPB>
F:313/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match          49.5%; Score 48.5; DB 1; Length 410;

```

```

Best Local Similarity 62.5%; Pred. No. 6;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY      4 GPTLRRLTSTRTPH 18
      ||:|||||:|:|:|
Db      298 GPTLRRLTSTRTPH 313

RESULT 3
C83365
2-oxoisovalerate dehydrogenase (alpha subunit) PA2247 [imported] - Pseudomonas aeruginos
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: C83365
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83365
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-410 <STO>
A:Cross-references: UNIPROT:Q91IM2; GB:AE004650; GB:AE004091; NID:99948267; PIDN:AAG0563
A:Experimental source: strain PA01
C:Genetics:
A:Gene: bkdA1; PA2247
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin

Query Match          49.5%; Score 48.5; DB 2; Length 410;
Best Local Similarity 62.5%; Pred. No. 6;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY      4 GPTLRRLTSTRTPH 18
      ||:|||||:|:|:|
Db      298 GPTLRRLTSTRTPH 313

RESULT 4
C75456
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: C75456
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
., M.; Shen, M.; Vanatrevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75456
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <WHI>
A:Cross-references: UNIPROT:Q9RV54; GB:AE001947; GB:AE000513; NID:96458665; PIDN:AAF1053
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0948
A:Map position: 1

Query Match          49.0%; Score 48; DB 2; Length 150;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      2 IEGLTRELTLSTRTPH 17
      ||:|||||:|:|:|
Db      97 LDGSLRLTSTRTPH 112

RESULT 5
G89894
protein kinase [imported] - Staphylococcus aureus (strain N315)

```





Db 237 EGPTLERTSYRTPHS 253

RESULT 9  
B64633  
hypothetical protein HP0906 - *Helicobacter pylori* (strain 26695)

C/Species: *Helicobacter pylori*  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C/Accession: B64633  
R/Tomb, J.F.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKernan, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.  
A/Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: B64633  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-527 <TOM>  
A/Cross-references: UNIPROT:O25564; GB:AE000600; GB:AE000511; NID:G2314042; PIDN:AAD0795

Query Match 47.4%; Score 46.5; DB 2; Length 527;  
Best Local Similarity 44.4%; Pred. No. 17;  
Matches 8; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 3 EGPTLREWLTSR--TPH 17  
:|||||: :  
98 QAPTLKCMINHKKTTPH 115

RESULT 10  
H83962  
serine/threonine protein kinase BH2504 [imported] - *Bacillus halodurans* (strain C-125)  
C/Species: *Bacillus halodurans*  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C/Accession: H83962  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: H83962  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-664 <STO>  
A/Cross-references: UNIPROT:Q9X920; GB:AP001515; GB:BA000004; NID:G10174886; PIDN:BA062  
A/Experimental source: strain C-125  
C/Genetics:  
A/Gene: BH2504

Query Match 46.9%; Score 46; DB 2; Length 664;  
Best Local Similarity 53.3%; Pred. No. 26;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IEGPTLREWLTSRTP 16  
:|||||: :  
Db 90 VEGPTLKEKIQQRKP 104

RESULT 11  
A59253  
myosin I beta - human  
C/Species: *Homo sapiens* (man)  
C/Date: 15-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C/Accession: A59253  
R/Crozat, F.; Amraoui, A.E.; Blanchard, S.; Lenoir, M.; Ripoll, C.; Vago, P.; Hamel, C.;  
Genomics 40, 332-341, 1997  
A/Title: Cloning of the genes encoding two murine and human cochlear unconventional type  
A/Reference number: A59253; MUID:97237053; PMID:9119401  
A/Accession: A59253  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA

A/Residues: 1-1028 <CRO>  
A/Cross-references: UNIPROT:O00159; GB:X98507; NID:G1926310; PIDN:CAA67131.1; PID:G19263  
A/Experimental source: dev stage adult; tissue type kidney  
C/Genetics:  
A/Gene: myo-1b  
A/Map position: 17p3.2-p13.3  
C/Superfamily: brush border myosin heavy chain I; myosin motor domain homology  
F.14-683/Domain: myosin motor domain homology <MMO>

Query Match 46.9%; Score 46; DB 2; Length 1028;  
Best Local Similarity 71.4%; Pred. No. 42;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SIEGPTLREWLTSR 14  
:|||||: :  
Db 308 SVEGPTLREALTHR 321

RESULT 12  
AD1246  
branched-chain alpha-keto acid dehydrogenase E1 chain (2-oxoisovalerate dehydrogenase a1  
C/Species: *Listeria monocytogenes*  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AD1246  
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A/Title: Comparative genomics of *Listeria* species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AD1246  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-331 <GLA>  
A/Cross-references: UNIPROT:O8Y7B4; GB:NC\_003210; PIDN:CAC99450.1; PID:G16410788; GSPDB:  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: lmoJ372  
C/Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain, thiamin pyrophosphate-bin

Query Match 46.4%; Score 45.5; DB 2; Length 331;  
Best Local Similarity 64.7%; Pred. No. 14;  
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 EGPTLREWLTSR-TPHS 18  
:|||||: :  
Db 234 EGPTLERTSYRTPHS 250

RESULT 13  
AH1608  
branched-chain alpha-keto acid dehydrogenase E1 chain (2-oxoisovalerate dehydrogenase a1  
C/Species: *Listeria innocua*  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AH1608  
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A/Title: Comparative genomics of *Listeria* species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AH1608  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-331 <GLA>  
A/Cross-references: UNIPROT:Q928Y3; GB:AL592022; PIDN:CMC96640.1; PID:G16413882; GSPDB:G  
A/Experimental source: strain C1p11262  
C/Genetics:  
A/Gene: lln1409

C/Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bi

Query Match 46.4%; Score 45.5; DB 2; Length 331;

Best Local Similarity 64.7%; Pred. No. 14;

Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 EGPFLREWLRS-TPHS 18  
 |||||:|||||  
 Db 234 EGPFLREWLRSYRTPHS 250

RESULT 14

S15618  
 E4 protein - human papillomavirus type 2a

C/Species: human papillomavirus type 2a

A/Note: host Homo sapiens (man)

C/Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004

C/Accession: S15618

R/Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.

Virus Res. 18, 81-98, 1990

A/Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and

A/Reference number: S15614; PMID:9186899; PMID:1964523

A/Accession: S15618

A/Molecule type: DNA

A/Residues: 1132 <HR>

A/Cross-references: UNIPROT:P25483; EMBL:X55964

C/Superfamily: papillomavirus type 2 E4 protein

C/Keywords: early protein

Query Match 45.9%; Score 45; DB 1; Length 132;

Best Local Similarity 40.6%; Pred. No. 6.2;

Matches 13; Conservative 2; Mismatches 1; Indels 16; Gaps 2;

QY 1 SIEGPTLR-----W-----LTSRTP 16  
 |||||:|||||  
 Db 90 SIEGPTLRSEKMGKSVTTSGLVTLTAQTP 121

RESULT 15

E84853  
 hypochlorite protein At2g42400 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: E84853

R/Lin, X.; Kall, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffet, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

eues, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; PMID:20083487; PMID:10617197

A/Accession: E84853

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1473 <SIO>

C/Accession: A32758

R/Rosen, D.R.; Martin-Morris, L.; Luo, L.; White, K.

Proc. Natl. Acad. Sci. U.S.A. 86, 2478-2482, 1989

A/Title: A Drosophila gene encoding a protein resembling the human beta-amyloid protein

A/Reference number: A32758; PMID:89184650; PMID:2494667

A/Accession: A32758

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1886 <ROS>

A/Cross-references: UNIPROT:P14599; GB:J04516; NID:q158371; PID:q158372

C/Genetics:

A/Genes: P14599

A/Cross-references: P14599

C/Keywords: transmembrane protein

Query Match 45.9%; Score 45; DB 2; Length 886;

Best Local Similarity 46.7%; Pred. No. 51;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTLRRLTSRTPHS 18  
 |||||:|||||  
 Db 829 GVAVAKRTSRSPHA 843

RESULT 17

S22383  
 axonin 1 precursor - chicken

N/Alternate names: neural cell adhesion molecule AxCAM

C/Species: Gallus gallus (chicken)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C/Accession: S22383; S34107; S69332; S22128

R/Zuellig, R.A.; Rader, C.; Schroeder, A.; Kalousek, M.B.; von Bohlen und Halbach, F.;

Bur, J. Biochem. 204, 453-463, 1992

A/Title: The axonally secreted cell adhesion molecule, axonin-1. Primary structure, imm

A/Reference number: S22383; PMID:92174898; PMID:1111675

A/Accession: S22383

A/Molecule type: mRNA

A/Residues: 1-1036 <ZUPL>

A/Cross-references: UNIPROT:P28685; EMBL:X63101; NID:q62852; PIDN:CAA44815.1; PID:q6285

A/Accession: S34107

A/Molecule type: protein

A/Residues: 29-49;51-80;84-95;100-117;120-128;130-141;143-176;243-254;256-296;303-336;3

R/Giger, R.U.; Vogt, L.; Zuellig, R.A.; Rader, C.; Henenhan-Beatty, A.; Wolfer, D.P.; So

Bur, J. Biochem. 227, 617-628, 1995

A/Title: The gene of chicken axonin-1. Complete structure and analysis of the promoter.

A/Reference number: S69332; PMID:95172044; PMID:7867620

A/Accession: S69332

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1002-1036 <GIG>

A/Cross-references: EMBL:X79607

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994

C/Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology

C/Keywords: cell adhesion

F.1.23/Domain: signal sequence #status predicted <SIG>

F.24-1036/Product: axonin 1 #status predicted <MAT>

F.336-392/Domain: immunoglobulin homology <IMM>

Query Match 45.9%; Score 45; DB 2; Length 1036;

Best Local Similarity 29.4%; Pred. No. 61;

Matches 10; Conservative 3; Mismatches 1; Indels 20; Gaps 1;

QY 5 PTLRE-----WTSRTPHS 18  
 |||||:|||||  
 Db 728 PTLRDYONGDFGYILSPKKGTCGWLTVRVPHA 761

RESULT 18

T13288  
 mel-41 protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004

C/Accession: T13288

R.Hart, K.L.; Santerre, A.; Sekelesky, J.J.; McKim, K.S.; Boyd, J.B.; Hawley, R.S.  
 Cell 82, 815-821, 1995  
 A:Title: The mei-41 gene of *D. melanogaster* is a structural and functional homolog of th  
 A:Reference number: Z11072; MUID:95401271; PMID:7671309  
 A:Accession: T13288  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2354 <HAR>  
 A:Cross-references: UNIPROT:Q24135; EMBL:U34925; NID:g998351; PID:g998351; PIDN:AA046881  
 C:Genetics:  
 A:Gene: mei-41  
 A:Cross-references: FlyBase:FBgn0004367  
 A:Introns: 650/3; 748/3; 2313/3  
 C:Function:  
 A:Description: involved in cell cycle checkpoint and meiotic recombination

Query Match 45.9%; Score 45; DB 2; Length 2354;  
 Best Local Similarity 56.2%; Pred. No. 1.5e+02;  
 Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 5 PTLREWLTSR--TPHS 18  
 | : ||| | ||| |  
 Db 2159 PVFOEWLRORFAPPHS 2174

RESULT 19  
 T23485  
 hypothetical protein K08F4.11 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T23485  
 R:Hembry, C.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: Z19746  
 A:Accession: T23485  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-200 <WIL>  
 A:Cross-references: EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN00022; CESP:K08F4.11  
 A:Experimental source: clone K08F4  
 C:Genetics:  
 A:Gene: CESP:K08F4.11  
 A:Map position: 4  
 A:Introns: 45/1; 76/1; 111/3  
 C:Superfamily: glutathione transferase

Query Match 44.9%; Score 44; DB 2; Length 200;  
 Best Local Similarity 61.5%; Pred. No. 14;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 IEQPTREWLTSR 14  
 ||| ||| ||| |  
 Db 183 IETPKLEWLAKR 195

RESULT 20  
 T37464  
 probable glutathione transferase (EC 2.5.1.18) GST3 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T37464  
 R:Rawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehren, K.  
 submitted to the EMBL Data Library, June 1997  
 A:Description: Parquat mediates differential gene expression in *C. elegans*.  
 A:Reference number: Z21702  
 A:Accession: T37464  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-207 <TRM>  
 A:Cross-references: UNIPROT:O16116; EMBL:AF010241; PIDN:AA054419.1  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: GST3

C:Superfamily: glutathione transferase  
 C:Keywords: transferase

Query Match 44.9%; Score 44; DB 2; Length 207;  
 Best Local Similarity 61.5%; Pred. No. 15;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 IEQPTREWLTSR 14  
 ||| ||| ||| |  
 Db 190 IETPKLEWLAKR 202

RESULT 21  
 C90277  
 hypothetical protein tmoA [imported] - *Sulfolobus solfataricus*  
 C:Species: *Sulfolobus solfataricus*  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C:Accession: C90277  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-  
 Jongs, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
 aretti, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: *Sulfolobus solfataricus* complete genome.  
 A:Reference number: A99139  
 A:Accession: C90277  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-310 <KID>  
 A:Cross-references: UNIPROT:Q97YT0; GB:AE006641; NID:g13814426; PIDN:AA041474.1; GSPDB:G  
 C:Genetics:  
 A:Gene: tmoA

Query Match 44.9%; Score 44; DB 2; Length 310;  
 Best Local Similarity 47.4%; Pred. No. 23;  
 Matches 9; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 SIEGPT-LREWLTSRTP 17  
 | ||| | : ||| : |||  
 Db 141 SFRGPTPDERKWLNEKYPH 159

RESULT 22  
 C87021  
 serine-threonine protein kinase [imported] - *Mycobacterium leprae*  
 C:Species: *Mycobacterium leprae*  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: C87021  
 R:Coile, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: C87021  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-400 <STO>  
 A:Cross-references: UNIPROT:O69568; GB:AL450380; NID:g13092968; PIDN:CAC31278.1; GSPDB:G  
 C:Genetics:  
 A:Gene: M0897  
 C:Superfamily: *Mycobacterium tuberculosis* probable serine/threonine-specific protein kin

Query Match 44.9%; Score 44; DB 2; Length 400;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEQPTREWLTSRTP 16  
 ||| ||| ||| |  
 Db 86 IEQPTREWLTSRTP 100

RESULT 23

Probable transposase VC0817 [similarity] - Vibrio cholerae (strain N16961 serogroup O1)  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C:Accession: E82277; T09435  
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qiu, H.; Dragol, I.; Sellers, F.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; PMID:20406833; PMID:10952301  
 A:Accession: E82277  
 A:Molecule type: DNA  
 A:Residues: 1-327 <HR>  
 A:Cross-references: UNIPROT:Q9KTS1; GB:AE004166; GB:AE003852; NID:g9655259; PIDN:AAF9398  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 R:Karrollis, D.K.R.; Johnson, J.A.; Bailey, C.C.; Boedeker, E.C.; Kapur, J.B.; Reeves, P.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 3134-3139, 1998  
 A:Title: A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains  
 A:Reference number: Z16672; PMID:9816509; PMID:9501228  
 A:Accession: T09435  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 3-327 <HR>  
 A:Cross-references: EMBL:AF034434; NID:g3004923; PIDN:AAI2272.1; PID:g3004924  
 A:Experimental source: strain N16961  
 C:Genetics:  
 A:Gene: VC0817  
 A:Map position: 1  
 A:Note: part of the pathogenicity island (VPI); associated with epidemic and pandemic strains

Query Match 44.4%; Score 43.5; DB 2; Length 327;  
 Best Local Similarity 36.7%; Pred. No. 29;  
 Matches 11; Conservative 1; Mismatches 5; Indels 13; Gaps 1;

QY 2 EGPTRLREWL-----TSRRPHS 18  
 DB 32 ISRPTLRKWKRRYKQCGIAGLESQSRPHS 61

RESULT 24  
 T06371  
 Probable UDP-glucuronosyltransferase (EC 2.4.1.-) - garden pea  
 C:Species: Pisum sativum (garden pea)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T06371  
 R:Woo, H.H.; Orbach, M.J.; Hawes, M.C.  
 submitted to the EMBL Data Library, November 1997  
 A:Description: lethal effects on root development by genetic alteration of glucuronide transferase  
 A:Reference number: Z15633  
 A:Accession: T06371  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-347 <MO>  
 A:Cross-references: UNIPROT:Q8J5C7; EMBL:AF034743; NID:g2827991; PIDN:AA899550.1; PID:g2827991  
 A:Experimental source: cv. Little Marvel  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 43.9%; Score 43; DB 2; Length 347;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 EGPTRLREWLTSRRPHS 18  
 DB 149 EEPCELEMLNSKEPNS 164

RESULT 25  
 B70936  
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) 2 - Mycobacterium tuberculosis  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: B70936

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:98295987; PMID:964220

A:Accession: B70936

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-399 <COL>

A:Cross-references: UNIPROT:O53510; GB:AL021957; GB:AL123456; NID:g3242293; PIDN:CAA1748

A:Experimental source: strain H37RV

C:Genetics:

C:Gene: pknJ

C:Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific protein kinase

C:Keywords: phosphotransferase

F:17-270/Domain: protein kinase homology <KIN>

Qy Query Match 43.9%; Score 43; DB 1; Length 399;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IEGPLREWLTSRT 16  
Db 99 IEGLTLELIERGP 113

RESULT 26

C98275

polyketide synthase and peptide synthetase (AB032549) [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C:Accession: C98275

R: Goodner, B.; Hinkle, C.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, S.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: C98275

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <KUR>

A:Cross-references: UNIPROT:Q8U9P9; GB:AE007870; PIDN:AAK69725.1; PID:g15159639; GSPDB:C98275

C:Genetics:

A:Gene: AGR\_L 2319

A:Map position: linear chromosome

C:Superfamily: ornithine-oxo-acid aminotransferase

Qy Query Match 43.9%; Score 43; DB 2; Length 491;  
Best Local Similarity 61.5%; Pred. No. 55;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EGPLREWLTSRT 15  
Db 383 DEPTLOEGLNART 395

RESULT 27

AC3009

polyketide synthase and peptide synthetase mcyE [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C:Accession: AC3009

R: Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan, S. Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Ew, A. A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AC3009

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-491 <KUR>  
A:Cross-references: UNIPROT:Q8U099, GB:AE008689, PIDN:AAL44489.1, PID:gl7742095, GSPDB:C  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: mcyE  
A:Map position: linear chromosome  
C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 43.9%; Score 43; DB 2; Length 491;  
Best Local Similarity 61.5%; Pred. No. 55;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 EGPTRWLTSTRTP 15  
:|||||:|:|  
Db 383 DGPTRLOEGINART 395

RESULT 28  
T44602  
phosphonate monoester hydrolase (EC 3.1.3.-) PEH [validated] - Burkholderia caryophylli  
C:Species: Burkholderia caryophylli  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T44602  
R:Dotson, S.B.; Smith, C.E.; Ling, C.S.; Barry, G.F.; Kishore, G.M.  
J. Biol. Chem. 271, 25754-25761, 1996  
A:Title: Identification, characterization and cloning of a phosphonate monoester hydrolase  
A:Reference number: Z22807; MUID:96421555; PMID:8824203  
A:Accession: T44602  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-514 <DOT>  
A:Cross-references: UNIPROT:Q45087, EMBL:U44852, NID:gl177863, PIDN:AA04467.1, PID:gl17  
A:Experimental source: strain PG2982  
A:Note: part of this sequence, including the amino end of the mature protein was confirm  
C:Genetics:  
A:Gene: pehA  
C:Complex: homotrimer [validated, MUID:96421555]  
C:Function:  
A:Description: EC 3.1.3.- [validated, MUID:96421555]; phosphonate monoester hydrolase; h  
A:Pathway: glyceryl phosphate utilization  
A:Note: may also function in vivo as phosphodiesterase  
C:Superfamily: animal sulfatase  
C:Keywords: homotrimer; phosphoric monoester hydrolase  
F:1-514/Product: phosphonate monoester hydrolase #status experimental <MAT>

Query Match 43.9%; Score 43; DB 2; Length 514;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 PTLREWLTSTRTPS 18  
||:|||||:  
Db 375 PTLREWLTSTRTPRA 388

RESULT 29  
H69878  
probable protein kinase (EC 2.7.1.-) yloP - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: H69878  
R:Kumar, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Beret  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Enlrich, S.D.; Emerson, P.T.; Entian, K.D.; Erttington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, S.; Galizzi, A.; Gallier  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
koetter, F.; Koningslet, G.; Krogh, S.; Kumano, M.; Kurita, K.; Kapitus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, R.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zamaite, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A65560; MUID:96044033; PMID:9364377  
A:Accession: H69878  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-648 <KUN>  
A:Cross-references: UNIPROT:Q34507, GB:Z29112, GB:AL009126, NID:G2633902, PIDN:CAB13450.  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yloP  
C:Superfamily: Bacillus subtilis probable protein kinase yloP; protein kinase homology  
C:Keywords: ATP; phosphotransferase; protein kinase  
F:9-269/Domain: protein kinase homology <KIN>

Query Match 43.9%; Score 43; DB 1; Length 648;  
Best Local Similarity 46.7%; Pred. No. 75;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IEGPTREWLTSTRTP 16  
:|||||:|:|  
Db 91 VEGMTLKEYITRANGP 105

RESULT 30  
F70697  
probable arabinosyltransferase - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: F70697  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: F70697  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1094 <COL>  
A:Cross-references: UNIPROT:P72060, GB:Z80343, GB:AL123456, NID:G3261648, PIDN:CAB02473.  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: emdA  
C:Superfamily: probable arabinosyl transferase

Query Match 43.9%; Score 43; DB 2; Length 1094;  
Best Local Similarity 61.5%; Pred. No. 1.4e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTREWLTSTRTP 16  
|||||:|:|  
Db 230 GRTLRDWLTSTRYP 242

RESULT 31  
T30939  
hemocyanin G-type chain - giant octopus (fragment)  
C:Species: Octopus dofleini (giant octopus)  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T30939  
R:Miller, K.I.; Culf, M.E.; Lang, W.F.; Varga-Weisz, P.; Field, K.G.; van Holde, K.E.  
J. Mol. Biol. 278, 827-841, 1998  
A:Title: Sequence of the Octopus dofleini hemocyanin subunit: structural and evolutionary  
A:Reference number: Z20940; MUID:98277150; PMID:9614945  
A:Accession: T30939  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2896 <MLI>  
A:Cross-references: UNIPROT:O61363, EMBL:AF020548, NID:G3132879, PID:G3132880, PIDN:AA03



C:Genetics:  
A:Note: Odcy  
C:Superfamily: hemocyanin

Query Match 43.4%; Score 43; DB 2; Length 2896;  
Best Local Similarity 53.8%; Pred. No. 4e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 TLRMTLSTRTPS 18  
DB 1427 TIRSWTGRDPHS 1439

## RESULT 32

DEBSPA

pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain - Bacillus subtilis

N:Alternate names: pyruvate dehydrogenase complex, E1 component alpha chain

C:Species: Bacillus subtilis

C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C:Accession: B36718; H68673

R:Hemilae, H.; Palva, A.; Paulin, L.; Arvidson, S.; Palva, I.

J. Bacteriol. 172, 5052-5063, 1990

A>Title: Secretary S complex of Bacillus subtilis: sequence analysis and identity to pyruvate dehydrogenase

A:Reference number: A36718; PMID:9036858; PMID:1697575

A:Accession: B36718

A:Molecule type: DNA

A:Residues: 1-371 <HEM>

A:Cross-references: UNIPROT:P21881; GB:M57435; GB:M31542; NID:g143375; PIDN:AAA62681.1;

R:Kunert, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berthel, C.; Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningslehn, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.;

Authors: Lauber, K.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli,

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, I.; Scanlon,

A:Authors: Schleich, S.; Schreier, R.; Scottone, P.; Sekiguchi, J.; Sekowska, A.; Seron,

akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; PMID:9804403; PMID:9384377

A:Accession: H69673

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1178 'A', 180-371 <KIN>

A:Cross-references: GB:299111; GB:AL009126; NID:g2633699; PIDN:CB13331.1; PID:g2633829

A:Experimental source: strain 168

C:Genetics:

A:Gene: pdhA

C:Function:

A>Description: catalyzes the decarboxylation of pyruvate coupled with formation of S-ace

A:Pathway: glycolysis

A:Note: uses thiamine diphosphate as a cofactor

C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bi

C:Keywords: flavoprotein; glycolysis; oxidoreductase; phosphoprotein; thiamin pyrophosph

F:2-371/Product: pyruvate dehydrogenase (lipoamide) alpha chain #status predicted <MAT>

F:165-212/Domain: thiamin pyrophosphate-binding domain homology <TPB>

Query Match 43.4%; Score 42.5; DB 1; Length 371;

Best Local Similarity 64.7%; Pred. No. 49;

Matches 11; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 3 EGPTLRMTLSR--TPHS 18

DB 259 EGPTLLETTRFRGPH 275

## RESULT 33

JC4742

transposase - Corynebacterium glutamicum

C:Species: Corynebacterium glutamicum

C>Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004

C:Accession: JC4742

R:Correia, A.; Pisabarro, A.; Castro, J.M.; Martin, J.F.

Gene 170, 91-94, 1996

A>Title: Cloning and characterization of an IS-like element present in the genome of Br

A:Reference number: JC4742; PMID:96200862; PMID:8621097

A:Accession: JC4742

A:Molecule type: DNA

A:Residues: 1-436 <COR>

A:Cross-references: UNIPROT:Q45293; EMBL:266534

A:Experimental source: ATCC 13869

A:Note: The authors translated the initiation codon TGT for residue 1 as Val

A:Note: The authors translated the codon ATT for residue 125 as Tyr

A:Note: the source is designated as Brevibacterium lactofermentum

C:Genetics:

A:Gene: GTG

F:388-415/Domain: DNA binding #status predicted <DNA>

F:405-415/Region: helix-turn-helix

Query Match 43.4%; Score 42.5; DB 2; Length 436;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 9; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 2 IEG---PTLRMTLSRTP 16  
DB 206 VEGRSADALRTMLAATP 223

## RESULT 34

B84271

glutamyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: B84271

R:Mc, M.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky,

; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl,

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L

A>Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; PMID:20504483; PMID:11016950

A:Accession: B84271

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-586 <STO>

A:Cross-references: UNIPROT:Q9HQ11; GB:AE004437; NID:g10580690; PIDN:AA619534.1; GSPDB:

C:Genetics:

A:Gene: glts

C:Superfamily: glutamine-tRNA ligase; glutamine-tRNA ligase homology

Query Match 43.4%; Score 42.5; DB 2; Length 586;  
Best Local Similarity 44.4%; Pred. No. 81;  
Matches 8; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 3 EGPTLRMTLSR--TPH 17  
DB 259 KNPALRDWAFRMDVDPH 276

## RESULT 35

AH1884

hypothetical protein al10625 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C:Accession: AH1884

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Matanabe, A.; Iriuch

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AH1884

A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-236 <KUR>  
 A:Cross-references: UNIPROT:Q8Y262, GB:BA000019, PIDN:BA072583.1, PID:917129971, GSPDB:C  
 A:Experimental source: strain PCC 7120  
 A:Genetics:  
 A:Gene: all0625

Query Match 42.9%; Score 42; DB 2; Length 236;  
 Best Local Similarity 58.3%; Pred. No. 35;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLEWLTGSRTP 16  
 | | | | | | | | | |  
 Db 60 PYLGEMLTNHP 71

RESULT 36  
 G87394  
 hypothetical protein CCL171 (imported) - Caulobacter crescentus

C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: G87394  
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwim, M.L.; Haft, D.H.; Kojon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87394  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-253 <STO>  
 A:Cross-references: UNIPROT:Q9A923, GB:AE005673, NID:g13422493, PIDN:AAK23155.1; GSPDB:C  
 A:Genetics:  
 A:Gene: CCL171  
 C:Superfamily: HCCA isomerase/mitochondrial glutathione S-transferase

Query Match 42.9%; Score 42; DB 2; Length 253;  
 Best Local Similarity 42.9%; Pred. No. 38;  
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLEWLTGSRTPS 18  
 | | | | | | | | | |  
 Db 53 PTLEWLTGSRTPS 66

RESULT 37

D70601  
 UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacteri  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: D70601

R:Colo, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 J.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: D70601

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-306 <COL>  
 A:Cross-references: UNIPROT:O05576, GB:Z94752, GB:AL123456, NID:g3261731, PIDN:CAB08153.  
 A:Experimental source: strain H37RV

C:Genetics:  
 A:Gene: galU  
 C:Superfamily: Escherichia coli UTP-glucose-1-phosphate uridylyltransferase  
 C:Keywords: nucleotidyltransferase

Query Match 42.9%; Score 42; DB 2; Length 306;  
 Best Local Similarity 63.6%; Pred. No. 47;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 GPTLEWLTGSR 14  
 | | | | | | | | | |  
 Db 290 GPDLLRRLLVAR 300

RESULT 38

U0119  
 Fc gamma (IgG) receptor IIb precursor - human  
 N:Alternate names: Fc gamma (IgG) receptor II (low affinity) beta; surface glycoprotein  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: U0119; A43543; A60568; A45877; S00478  
 R:Brooks, D.G.; Qiu, W.O.; Luster, A.D.; Ravetch, J.V.  
 J. Exp. Med. 170, 1369-1385, 1989

A:Title: Structure and expression of human IgG FcRRII (CD32): functional heterogeneity is  
 A:Reference number: U0118; MUID:90010791; PMID:2529342  
 A:Accession: U0119

A:Molecule type: mRNA  
 A:Residues: 1-310 <BRO>

A:Cross-references: UNIPROT:P31994  
 R:Engelhardt, W.; Geerdts, C.; Frey, J.  
 Eur. J. Immunol. 20, 1367-1377, 1990

A:Title: Distribution, inducibility and biological function of the cloned and expressed  
 A:Reference number: A43543; MUID:90316181; PMID:2142460  
 A:Accession: A43543

A:Molecule type: mRNA

A:Residues: 1-204, 'Y', 206-254, 274-310 <ENG>  
 A:Cross-references: GB:X52473; NID:g3928171; PIDN:CAA36713.1; PID:g29428

R:Engelhardt, W.; Geerdts, C.; Frey, J.  
 Mol. Immunol. 27, 379-382, 1990

A:Title: Organization of human FCRII and FCRII-like (betaFCRII) genes: structural homolo  
 A:Reference number: A60568; MUID:90294837; PMID:2141667

A:Accession: A60568

A:Molecule type: DNA  
 A:Residues: 1-38 <EN2>

R:Seiki, T.

Immunogenetics 30, 5-12, 1989  
 A:Title: Identification of multiple isoforms of the low-affinity human IgG Fc receptor.

A:Reference number: A45877; MUID:89307398; PMID:2526077

A:Accession: A45877

A:Status: preliminary

A:Molecule type: mRNA  
 A:Residues: 1-74, 'O', '76-119', 'V', 121-204, 'Y', 206-253, 'T', 233-254, 274-310 <SEK>

A:Cross-references: GB:M28696; NID:g184843; PIDN:AA36051.1; PID:g306929  
 A:Note: the authors translated the codon CAG for residue 75 as His  
 R:Stengelin, S.; Stamenkovic, I.; Seed, B.  
 EMBO J. 7, 1053-1059, 1988

A:Title: Isolation of cDNAs for two distinct human Fc receptors by ligand affinity cloni  
 A:Reference number: S00477; MUID:88296409; PMID:3402431

A:Accession: S00478  
 A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-35, 'S', '37-204', 'Y', 206-253, 'G', 255 <STB>  
 A:Note: the authors suggest that the cDNA is derived from a precursor RNA that still con

C:Genetics:

A:Gene: GDB:FCGR2B; FCG2; FCGR2  
 A:Cross-references: GDB:18183; OMIM:146790

A:Map position: 1q23-1q23  
 A:Introns: 131/1

C:Superfamily: Fc gamma receptor II; immunoglobulin homology  
 C:Keywords: alternative splicing; glycoprotein; immunoglobulin, immunoglobulin receptor,

F:1-44/Domain: signal sequence #status predicted <STG>  
 F:45-310/Product: IgG Fc receptor IIb #status predicted <MAT>

F:64-115/Domain: extracellular #status predicted <EXT>  
 F:145-198/Domain: immunoglobulin homology <IMM2>

F:223-245/Domain: transmembrane #status predicted <INT>  
 F:246-310/Domain: intracellular #status predicted <INT>

F:106,180,187/Binding site: carbohydrate (Asn (covalent) #status predicted

Query Match 42.9%; Score 42; DB 2; Length 310;



Best Local Similarity 63.6%; Pred. No. 48;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LREWLTSRTPH 17  
| | | | |  
: : : : :  
Db 129 LSEWLVLTQTPH 139

## RESULT 39

JL0118  
Fc gamma (IgG) receptor IIA precursor - human  
N:Alternate names: Fc gamma (IgG) receptor II (low affinity) alpha; surface glycoprotein  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C/Accession: JL0118; A31932; S02297; B45877; S00477; S02296  
R:Brooks, D.G.; Qiu, W.Q.; Lunster, A.D.; Ravetch, J.V.  
J. Exp. Med. 170, 1369-1385, 1989  
A>Title: Structure and expression of human IgG FcRII (CD32): functional heterogeneity is  
A/Reference number: JL0118; MUID:90010791; PMID:2529342  
A/Accession: JL0118  
A/Molecule type: mRNA  
A/Residues: 1-317 <PRO>  
A/Cross-references: UNIPROT:P12318; GB:M31932; NID:g182473; PID:AAA35827.1; PID:g182474  
A:Experimental source: placenta  
A>Note: It is uncertain whether Met-1, Met-3, or Met-7 is the initiator  
R:Hibbs, M.L.; Bonadonna, L.; Scott, B.M.; McKenzie, I.F.C.; Hogarth, P.M.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2240-2244, 1988  
A>Title: Molecular cloning of a human immunoglobulin G Fc receptor.  
A/Reference number: S02296; MUID:88176920; PMID:2965389  
A/Accession: A31932  
A/Molecule type: mRNA  
A/Residues: 3-317 <HIB>  
A/Cross-references: EMBL:J03619; NID:g183619; PID:AAA35932.1; PID:g306803  
R:Stuart, S.G.; Trounstein, M.L.; Vaux, D.J.T.; Koch, T.; Martens, C.L.; Mellman, I.; Mc  
J. Exp. Med. 166, 1668-1684, 1987  
A>Title: Isolation and expression of cDNA clones encoding a human receptor for IgG (Fc-gamma-1).  
A/Reference number: S02297; MUID:88061079; PMID:2824655  
A/Accession: S02297  
A/Molecule type: mRNA  
A/Residues: 1, 'T', 3-317 <STU>  
A/Cross-references: EMBL:Y00644; NID:g31335; PID:CAA6872.1; PID:g31336  
A>Note: It is uncertain whether Met-1 or Met-7 is the initiator  
R:Seiki, T  
Immunogenetics 30, 5-12, 1989  
A>Title: Identification of multiple isoforms of the low-affinity human IgG Fc receptor.  
A/Reference number: A45877; MUID:9307398; PMID:2526077  
A/Accession: B45877  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 7-317 <SEK>  
A/Cross-references: GB:M28697; NID:g184841; PID:AAA36050.1; PID:g306928  
R:Stengel, S.; Stamenkovic, I.; Seed, B.  
EMBO J. 7, 1053-1059, 1988  
A>Title: Isolation of cDNAs for two distinct human Fc receptors by ligand affinity cloning.  
A/Reference number: S00477; MUID:88296409; PMID:3402431  
A/Accession: S00477  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 7-317 <STB>  
C/Genetics:  
A:Gene: GDB:FCGR2A  
A/Cross-references: GDB:119903; OMIM:146790  
A/Map position: 1Q23-1Q23  
C:Superfamily: Fc gamma receptor III; immunoglobulin homology  
C/Keywords: glycoprotein; immunoglobulin receptor; transmembrane protein  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:36-317/Product: IgG Fc receptor IIA #status predicted <REI>  
F:36-216/Domain: extracellular #status predicted <EXT>  
F:55-106/Domain: immunoglobulin homology <IMM1>  
F:136-189/Domain: immunoglobulin homology <IMM2>  
F:217-240/Domain: transmembrane #status predicted <TM>  
F:241-317/Domain: intracellular #status predicted <INT>

F:97,178/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 42.9%; Score 42; DB 2; Length 317;  
Best Local Similarity 63.6%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LREWLTSRTPH 17  
| | | | |  
: : : : :  
Db 120 LSEWLVLTQTPH 130

## RESULT 40

S06946  
Fc gamma (IgG) receptor - human  
C:Species: Homo sapiens (man)  
C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C/Accession: S06946  
R:Stuart, S.G.; Simister, N.E.; Clarkson, S.B.; Kacinski, B.M.; Shapiro, M.; Mellman, I  
EMBO J. 8, 3657-3666, 1989  
A>Title: Human IgG Fc receptor (FcRII; CD32) exists as multiple isoforms in macrophage  
A/Reference number: S06946; MUID:90059965; PMID:2531080  
A/Accession: S06946  
A/Molecule type: mRNA  
A/Residues: 1-323 <STU>  
A/Cross-references: UNIPROT:P11995; EMBL:X17652; NID:g32073; PID:CAA35642.1; PID:g3207  
C:Superfamily: Fc gamma receptor III; immunoglobulin homology  
C/Keywords: immunoglobulin receptor; transmembrane protein  
F:64-115/Domain: immunoglobulin homology <IMM>

Query Match 42.9%; Score 42; DB 2; Length 323;  
Best Local Similarity 63.6%; Pred. No. 50;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LREWLTSRTPH 17  
| | | | |  
: : : : :  
Db 129 LSEWLVLTQTPH 139

## RESULT 41

B82096  
conserved hypothetical protein VC2278 [imported] - Vibrio cholerae (strain N16961 serog  
C:Species: Vibrio cholerae  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C/Accession: B82096  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragoi, I.; Sellers,  
1, R.R.; Mekalanos, J.D.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A/Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: B82096  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-430 <HEI>  
A/Cross-references: UNIPROT:Q9KPT4; GB:AE004299; GB:AE003852; NID:g9656835; PID:AAF954  
A/Experimental source: serogroup O1; strain N16961; biotype El Tor  
C/Genetics:  
A:Gene: VC2278  
A/Map position: 1  
C:Superfamily: conserved hypothetical protein HI0125

Query Match 42.9%; Score 42; DB 2; Length 430;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LREWLTSRTPHS 18  
: | | | | : | | | |  
: | | | | : | | | |  
Db 119 IREWLINSIPHS 130

## RESULT 42

B45438  
myosin I beta, MMI beta - mouse (fragment)

C/Species: Mus musculus (house mouse)  
C/Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: B45438  
R/Sherr, E.H.; Joyce, M.P.; Greene, L.A.  
J. Cell Biol. 120, 1405-1416, 1993  
A/Title: Mammalian myosin I alpha, I beta, and I gamma: new widely expressed genes of th  
A/Reference number: A45438; MUID:93194946; PMID:8449986  
A/Accession: B45438  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: nucleic acid  
A/Residues: 1-448 <SHE>  
A/Cross-references: UNIPROT:Q9WT17  
A/Note: sequence extracted from NCBI backbone (NCBI:P131911)  
C/Suprafamily: brush border myosin heavy chain I; myosin motor domain homology  
F1-448/Domain: myosin motor domain homology (fragment) <MMOT>  
  
Query Match 42.9%; Score 42; DB 2; Length 448;  
Best Local Similarity 69.2%; Pred. No. 72;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 IEGPTLRWLTSTR 14  
: ||||| |||  
DB 164 VEGTTLREALTLTR 176  
  
RESULT 43  
T37934  
conserved hypothetical protein SPAC1952.06c - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
R/McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.  
submitted to the EMBL Data Library, August 1999  
A/Reference number: Z21755  
A/Accession: T37934  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-564 <MKCD>  
A/Cross-references: UNIPROT:Q9UTK1; EMBL:AL109820; PIDN:CA852570.1; GSPDB:GN00066; SPDB:  
A/Experimental source: strain 972h-; cosmid c1952  
C/Genetics:  
A/Gene: SPDB:SPAC1952.06c  
A/Map position: 1  
  
Query Match 42.9%; Score 42; DB 2; Length 564;  
Best Local Similarity 57.1%; Pred. No. 93;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 2 IEGPTLRWLTSTR 15  
: ||||| |||  
DB 238 VENTTLVENLTLRS 251  
  
RESULT 44  
A39707  
erythrocyte membrane band 4.2 protein - human  
N/Alternate names: pallidin  
N/Contents: erythrocyte membrane band 4.2 protein, long splice form; erythrocyte membra  
C/Species: Homo sapiens (man)  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: A39707; A34865; B34865; A34883  
R/Korogren, C.; Cohen, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4840-4844, 1991  
A/Title: Organization of the gene for human erythrocyte membrane protein 4.2: structural  
A/Reference number: A39707; MUID:91271288; PMID:2052563  
A/Accession: A39707  
A/Molecule type: DNA  
A/Residues: 1-721 <ROR1>  
A/Cross-references: UNIPROT:P16452; GB:L06519; NID:g306738; PIDN:AA52385.1; PID:g306740  
A/Experimental source: cell type erythrocyte; tissue type peripheral blood; tissue lib h  
R/Sung, L.A.; Chien, S.; Chang, L.S.; Lambert, K.; Bliss, S.A.; Bouhassira, E.E.; Nagel,  
Proc. Natl. Acad. Sci. U.S.A. 87, 955-959, 1990  
A/Title: Molecular cloning of human protein 4.2: a major component of the erythrocyte me

A/Reference number: A34865; MUID:90138995; PMID:1689063  
A/Accession: A34865  
A/Molecule type: mRNA  
A/Residues: 1-364, 'KRGLPC', 371-379, 'H', 381-405, 'L', 407-721 <SUN1>  
A/Cross-references: GB:M0647; NID:g189433; PIDN:AAA36401.1; PID:g189434  
A/Accession: B34865  
A/Molecule type: mRNA  
A/Residues: 1-3, 34-364, 'KRGLPC', 371-379, 'H', 381-405, 'L', 407-721 <SUN2>  
A/Cross-references: GB:M0646; NID:g189435; PIDN:AAA36402.1; PID:g189436  
A/Experimental source: isolate Sickle cell patient; cell type reticulocyte  
A/Note: parts of this sequence were determined by protein sequencing  
R/Korogren, C.; Lawler, J.; Lambert, S.; Spechter, D.; Cohen, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 87, 613-617, 1990  
A/Title: Complete amino acid sequence and homologies of human erythrocyte membrane prote  
A/Reference number: A34883; MUID:90138879; PMID:2300550  
A/Accession: A34883  
A/Molecule type: mRNA  
A/Residues: 1-3, 34-721 <KOR2>  
A/Cross-references: GB:M29399; NID:g182083; PIDN:AAA35798.1; PID:g182084  
C/Comment: This protein is a major constituent of the erythrocyte membrane. It apparentl  
C/Genetics:  
A/Gene: GDB:EPB42; PA  
A/Cross-references: GDB:127385; OMIM:177070  
A/Map position: 15q15-15q15  
C/Suprafamily: protein-glutamine gamma-glutamyltransferase  
C/Keywords: alternative splicing; blocked amino end; glycoprotein; lipoprotein; myristyl  
F1-2-721/Product: erythrocyte membrane band 4.2 protein, long splice form #status predict  
F1-2-34-721/Product: erythrocyte membrane band 4.2 protein, short splice form #status p  
F1-298-316/Domain: transmembrane #status predicted <IRM>  
F1-518-520/Region: cell attachment (R-G-D) motif (in mature form) #status predicted  
F1-2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F1-103-420, 447, 529, 604, 705/Binding site: carbohydrate (asn) (covalent) #status predict  
F1-278/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predic  
  
Query Match 42.9%; Score 42; DB 2; Length 721;  
Best Local Similarity 70.0%; Pred. No. 12e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 PTLRWMLTSTR 14  
: ||||| |||  
DB 280 PTLRWMLTSTR 289  
  
RESULT 45  
H75634  
myosin-Ic - mouse (fragment)  
N/Alternate names: myosin-I beta  
C/Species: Mus musculus (house mouse)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: H75634  
R/Crozet, F.; Amraoui, A.E.; Blanchard, S.; Lenoir, M.; Ripoll, C.; Vago, P.; Hamel, C.;  
Genomics 40, 332-341, 1997  
A/Title: Cloning of the genes encoding two murine and human cochlear unconventional type  
A/Reference number: A59253; MUID:91237053; PMID:9119901  
A/Accession: H75634  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-807 <CRO>  
A/Cross-references: UNIPROT:Q9WT17; GB:X99638; NID:g1924960; PIDN:CA67956.1; PID:g19249  
A/Experimental source: strain BALB/c; tissue type cochlea; dev stage adult  
C/Genetics:  
A/Gene: MGI:Myo1c  
A/Cross-references: MGI:106612  
A/Map position: 11:44.1  
C/Suprafamily: brush border myosin heavy chain I; myosin motor domain homology  
F1-14-683/Domain: myosin motor domain homology <MMO>  
  
Query Match 42.9%; Score 42; DB 2; Length 807;  
Best Local Similarity 69.2%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 IEGPTLRWLTSTR 14  
: ||||| |||

Db 309 VEGTILREALTR 321

Search completed: September 1, 2005, 16:22:56  
Job time : 15.7266 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 66.9496 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-10

Perfect score: 98  
Sequence: 1 SIEGPTLRWLTSRTPHS 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	55.6	332	2	08CP43
2	52.5	53.6	410	2	062SV1
3	52.5	53.6	410	2	063H26
4	51	52.0	753	2	09YB62
5	51.0	52.0	295	2	06DG77
6	49	50.0	466	2	089B08
7	49	50.0	576	2	08UDD0
8	49	50.0	1049	2	09XBP6
9	48.5	49.5	410	1	08BA_PSEPU
10	48.5	49.5	410	2	088E02
11	48.5	49.5	410	2	091IM2
12	48	49.0	150	2	09RV54
13	48	49.0	245	2	066272
14	48	49.0	249	2	082989
15	48	49.0	278	2	09XDV0
16	48	49.0	319	2	09RKM5
17	48	49.0	388	2	09KX10
18	48	49.0	481	2	08GAT0
19	48	49.0	632	2	0751T9
20	48	49.0	664	2	08NX14
21	48	49.0	664	2	099UP8
22	48	49.0	664	2	07A5Z8
23	48	49.0	664	2	06G9Z3
24	48	49.0	664	2	06GHI5
25	47	48.0	91	2	08Y0T5
26	47	48.0	154	1	YD55_HABIN
27	47	48.0	245	2	066278
28	47	48.0	282	2	070BP7
29	47	48.0	417	2	06J6D5
30	47	48.0	430	2	031703
31	47	48.0	667	2	08CSV9

32	47	48.0	818	2	07WMY0	Q7WMY0	alcaligenes
33	46.5	47.4	330	1	0DBA_BACSU	P37940	bacillus su
34	46.5	47.4	330	2	065HK8	065HK8	bacillus li
35	46.5	47.4	527	2	025564	025564	helicobacte
36	46	46.9	156	2	064K11	064K11	eleutheroda
37	46	46.9	156	2	064K12	064K12	eleutheroda
38	46	46.9	156	2	064K16	064K16	eleutheroda
39	46	46.9	156	2	064K24	064K24	eleutheroda
40	46	46.9	156	2	064K25	064K25	eleutheroda
41	46	46.9	377	2	082PK5	082PK5	streptomyce
42	46	46.9	559	2	0745Z3	0745Z3	thermus the
43	46	46.9	570	1	SYE_PYPAB	08ZU33	pyrobaculum
44	46	46.9	664	2	09K920	09K920	bacillus ha
45	46	46.9	1028	1	MYIC_HUMAN	000159	homo sapien
46	46	46.9	1028	2	06NVJ7	06NVJ7	homo sapien
47	46	46.9	1030	2	086Y95	086Y95	homo sapien
48	45.5	46.4	168	2	09V492	09V492	desulfohal
49	45.5	46.4	292	2	072B10	072B10	desulfovibr
50	45.5	46.4	328	2	097317	097317	plasmodium
51	45.5	46.4	331	2	08Y7B4	08Y7B4	listeria mo
52	45.5	46.4	331	2	092BY3	092BY3	listeria in
53	45.5	46.4	331	2	071EVO	071EVO	listeria mo
54	45.5	46.4	332	2	08GLC9	08GLC9	listeria mo
55	45	45.9	132	1	VE4_HPVA	091C09	human papil
56	45	45.9	244	2	066269	066269	erythromicr
57	45	45.9	244	2	0987K1	0987K1	erythroba
58	45	45.9	245	2	082991	082991	erythroba
59	45	45.9	245	2	092N87	092N87	porphyrobac
60	45	45.9	246	2	066276	066276	porphyrobac
61	45	45.9	326	2	P95613	P95613	rhizobium g
62	45	45.9	444	2	08AZ50	08AZ50	infectious
63	45	45.9	444	2	08BDV4	08BDV4	infectious
64	45	45.9	450	2	09SLB9	09SLB9	arabidopsis
65	45	45.9	460	2	065854	065854	beet yellow
66	45	45.9	494	2	06NHQ1	06NHQ1	corynebacte
67	45	45.9	542	2	06APS4	06APS4	desulfocale
68	45	45.9	658	2	09IAC1	09IAC1	brachydanio
69	45	45.9	887	1	A4_DROME	P14599	drosofila
70	45	45.9	1036	1	AXO1_CHICK	P28655	gallus gall
71	45	45.9	2354	2	024135	P24135	drosofila
72	45	45.9	2517	2	09VXG9	09VXG9	drosofila
73	44.5	45.4	342	2	072K49	072K49	thermus the
74	44.5	45.4	661	2	096ZC0	096ZC0	caenorhabdi
75	44.5	45.4	669	2	096ND7	096ND7	caenorhabdi
76	44.5	45.4	1009	2	07SEY7	07SEY7	neutrospora
77	44	44.9	191	2	0938S9	0938S9	uncultured
78	44	44.9	207	1	GTS3_CABEL	016116	caenorhabdi
79	44	44.9	245	2	082987	082987	erythroba
80	44	44.9	252	2	08XP09	08XP09	ralstonia s
81	44	44.9	267	2	074FH6	074FH6	geobacter s
82	44	44.9	302	2	0742B3	0742B3	mycobacteri
83	44	44.9	309	2	073H24	073H24	wolbachia p
84	44	44.9	310	2	097YTO	097YTO	sulfolobum
85	44	44.9	354	2	08ZVT5	08ZVT5	pyrobaculum
86	44	44.9	375	2	07XBP6	07XBP6	oryza sativ
87	44	44.9	400	2	069568	069568	mycobacteri
88	44	44.9	492	1	Y193_COREF	08F922	corynebacte
89	44	44.9	941	2	08OUJ6	08QUJ6	infectious
90	44	44.9	984	2	06MQ52	06MQ52	bdellovibri
91	44	44.9	1028	2	028138	028138	bos taurus
92	44	44.9	5953	2	08XS39	08XS39	ralstonia s
93	44	44.9	6889	2	08XS40	08XS40	ralstonia s
94	43.5	44.4	216	2	09CR76	09CR76	m mus muscu
95	43.5	44.4	227	2	06A689	06A689	leifsonia x
96	43.5	44.4	325	2	068334	068334	vibrio chol
97	43.5	44.4	327	2	09KTS1	09KTS1	vibrio chol
98	43.5	44.4	340	2	07XIL8	07XIL8	oryza sativ
99	43	43.9	118	2	07VB61	07VB61	prochloroco
100	43	43.9	157	2	09NMG3	09NMG3	homo sapien

## ALIGNMENTS

```

RESULT 1
ID 08CP43 PRELIMINARY; PRT; 332 AA.
AC 08CP43;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Branched-chain alpha-keto acid dehydrogenase E1.
GN OrderedlocusNames=SE1198;
OC Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
  Qin Z.-Q., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.,
  Yuan Z.-H., Zhao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
  "Genome-based analysis of virulence genes in a non-biofilm-forming
  Staphylococcus epidermidis strain (ATCC 12228).";
RT Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AE016748; AA004797.1; -.
DR HSSP; P12694; 10LX.
DR GO; GO:0016224; F:oxidoreductase activity, acting on the alde. .; IEA.
DR GO; GO:0008152; P:metabolism, IEA.
DR InterPro; IPR001017; Dehydrogenase_E1.
DR Pfam; PF00676; E1_dh; 1.
DR Complete proteome.
SQ SEQUENCE 332 AA; 36682 MW; 0A9B4468EC96C975 CRC64;

```

```

Query Match 55.6%; Score 54.5; DB 2; Length 332;
Best Local Similarity 76.5%; Pred. No. 2.5;
Matches 13; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 3 EGPTLEWLTSTR-TPHS 18
Db 234 EGPTLEWLTSTRTPHS 250

```

```

RESULT 2
ID 0629V1 PRELIMINARY; PRT; 410 AA.
AC 0629V1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 2-oxoisovalerate dehydrogenase, E1 component, alpha subunit (EC
  1.2.4.4).
GN ORFNames=BMAA2013;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Nierman W.C., Deshaizer D., Kim H.S., Tettelin H., Nelson K.E.,
  Feldblum T., Ulrich R.L., Romling C.M., Brinkac L.M., Daugherty S.C.,
  Davidson T.D., Deboy R.T., Dmitrov G., Dodson R.J., Durkin A.S.,
  Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
  Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarría S.,
  Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafer N.,
  Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000011; ANU45601.1; -.
KW Oxidoreductase.
SQ SEQUENCE 410 AA; 45445 MW; 99DF85B96288CC2 CRC64;

```

```

Query Match 53.6%; Score 52.5; DB 2; Length 410;
Best Local Similarity 68.8%; Pred. No. 6.6;

```

```

Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 4 GPTLEWLTSTR-TPHS 18
Db 298 GPTLEWLTSTRTPHS 313

```

```

RESULT 3
ID 063H26 PRELIMINARY; PRT; 410 AA.
AC 063H26;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 2-oxoisovalerate dehydrogenase alpha subunit (EC 1.2.4.4).
GN Name=bkdA1; ORFNames=BPS2273;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Tlball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
  Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
  Bentley S.D., Sebahia M., Thomson N.R., Baeson N., Beacham I.,
  Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
  Chillingworth T., Cronin A., Crosset B., Davis P., Deshaizer D.,
  Felwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagers K.,
  Keith K.E., Maddison M., Moule S., Price C., Quail W.A.,
  Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
  Songvilay S., Stevens K., Tumapa S., Vesaratchavee M.,
  Whitehead S., Yeats C., Barrett B.G., Oyston P.C.F., Parthill J.;
RA "Genomic plasticity of the causative agent of melioidosis,
  Burkholderia pseudomallei";
RT Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH39759.1; -.
KW Oxidoreductase.
SQ SEQUENCE 410 AA; 45415 MW; 99C9C7DD52E88CC2 CRC64;

```

```

Query Match 53.6%; Score 52.5; DB 2; Length 410;
Best Local Similarity 68.8%; Pred. No. 6.6;
Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
QY 4 GPTLEWLTSTR-TPHS 18
Db 298 GPTLEWLTSTRTPHS 313

```

```

RESULT 4
ID 09YE62 PRELIMINARY; PRT; 753 AA.
AC 09YE62;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 753aa long hypothetical aldehyde oxidoreductase.
GN OrderedlocusNames=APB0708;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KL;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
  Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Koeugi H.,
  Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
  Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
  Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
  Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

```

Query Match	52.0%	Score 51	DB 2	Length 753
Best Local Similarity	47.1%	Pred. No. 22		
Matches 8	Conservative 5	Mismatches 4	Indels 0	Gaps 0
Db	1	SYDSTLTWVSTQTPH 208		
Cy	1	SIEGPTLAEMUTSRTPH 17		
RESULT 5				
06DG77				
AC	06DG77	PRELIMINARY	PRT	295 AA.
DT	25-OCT-2004 (TREMBLrel. 28, Created)			
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)			
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE	Zgc:91907.			
GN	Name=zgc:91907			
OS	Brachydanio rerio (zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxId=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Whole;			
RX	MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Gronow L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Murusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.O., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bohak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whitesley R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Kozminski A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzyszowski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Whole;			
RA	Strausberg R.;			
RA	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC076474; AAH76474.1; -			
DR	GO; GO:0046872; F:metal ion binding; IEA.			
DR	InterPro; IPR001594; Znf_DHHC.			
DR	Pfam; PF01529; zf_DHHC; 1			
DR	ProDom; PD003041; Znf_DHHC; 1.			
DR	PROSITE; PS50216; ZF_DHHC; 1.			

QY	DB	SEQUENCE	295 AA;	33706 MM;	A89FE24A30E6CEAF CRC64;	51.0%;	Score 50;	DB 2;	Length 295;	Best Local Similarity 47.1%;	Pred. No. 12;	Matches 8;	Conservative 4;	Mismatches 5;	Indels 0;	Gaps 0;
QY	230	IEGPTLRWLTSTRPYS 18														
DB	230	LRGQTTREWYSTRPYS 246														
RESULT 6																
Q89EQ8	PRELIMINARY;	PRT;	466 AA.													
AC	Q89EQ8;															
DT	01-JUN-2003 (TREMBLrel. 24, Created)															
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)															
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)															
DE	B117014 protein.															
GN	OrderedlocusNames=b117014;															
OS	Bradyrhizobium japonicum.															
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;															
OC	Bradyrhizobiaceae; Bradyrhizobium.															
OX	NCBI_TaxID=375;															
RN	[1]															
RP	SEQUENCE FROM N.A.															
RC	STRAIN=USD110;															
RX	MEDLINE=22484998; PubMed=12597275;															
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchimi T.,															
RA	Sasamoto S., Watanabe A., Ideawa K., Iriyuchi M., Kawashima K.,															
RA	Kohara M., Matsumoto M., Shingo S., Tsunooka H., Wada T., Yamada M.,															
RA	Tabata S.,															
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium															
RT	Bradyrhizobium japonicum USD110."															
RL	DNA Res. 9:189-197(2002).															
DR	EMBL; AP005960; BACS2279.1; "															
DR	GO; GO:0016747; P:transferase activity, transferring groups o. . .; IBA.															
DR	InterPro; IPR002656; Acyl_transf_3.															
DR	Pfam; PF01757; Acyl_transf_3; 1.															
KM	Complete proteome.															
SEQ	SEQUENCE 466 AA; 51100 MM; F5489F1904D701F9 CRC64;															
Query Match			50.0%;	Score 49;	DB 2;	Length 466;										
Best Local Similarity			64.3%;	Pred. No. 28;												
Matches			9;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;									
QY	2	IEGPTLRWLTSTRT 15														
DB	234	IEGKVRPWTSTRT 247														
RESULT 7																
Q8UOD0	PRELIMINARY;	PRT;	576 AA.													
AC	Q8UOD0;															
DT	01-JUN-2002 (TREMBLrel. 21, Created)															
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)															
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)															
DE	Hypothetical protein PF1669.															

```

Query Match      50.0%; Score 49; DB 2; Length 576;
Best Local Similarity 72.7%; Pred. No. 35;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      8 REMLTSRTPHS 18
      |||||
      158 REMLSTNTPHT 168

RESULT 8
ID Q9XBP6 PRELIMINARY; PRT; 1049 AA.
AC Q9XBP6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serine/threonine kinase PK08.
GN Name=pk08;
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytochrome b; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZF1;
RA Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,
RA Munoz-Dorado J., Farez-Vidal E., Inouye M.,
RA Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF159691; MAD42856.1; -.
DR HSSP; P71584; 1067.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR CO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR008940; Prey1_kinase.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR_1; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00028; TPR; 4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50005; TPR; 2.
DR PROSITE; PS50293; TPR_REGION; 1.
DR ATP-binding; Kinase; Repeat; TPR repeat.
DR ATP-binding; Kinase; Repeat; TPR repeat.
SQ SEQUENCE 1049 AA; 114312 MW; 7752862DAA25338C CRC64;

Query Match      50.0%; Score 49; DB 2; Length 1049;
Best Local Similarity 53.3%; Pred. No. 68;
Matches      8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      2 IEGPTLEWLTSTRTP 16
      |||||
      167 VEGTTLAEWKKERRP 181

RESULT 9
ID ODBA_PSEPU STANDARD; PRT; 410 AA.
AC P09060;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 2-oxoisovalerate dehydrogenase alpha subunit (EC 1.2.4.4) (Branched-
DE chain alpha-keto acid dehydrogenase E1 component alpha chain) (BCKDH
DE E1-alpha).
GN Name=bkdA1;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=103;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RX MEDLINE=88329084; PubMed=3416875;
RA Burns G., Brown T., Hatter K., Jarvis J., Sokatch J.R.;
RT "Similarity of the E1 subunits of branched-chain-oxoacid dehydrogenase
RT from Pseudomonas putida to the corresponding subunits of mammalian
RT branched-chain-oxoacid and pyruvate dehydrogenases.";
RL Eur. J. Biochem. 176:311-317(1988).
RN [2]
RP SEQUENCE OF 1-17 FROM N.A.
RC STRAIN=G2;
RX MEDLINE=9356017; PubMed=2211503;
RA Madhusudan K.T., Huang G., Burns G., Sokatch J.R.;
RT "Transcriptional analysis of the promoter region of the Pseudomonas
RT putida branched-chain keto acid dehydrogenase operon.";
RL J. Bacteriol. 172:5655-5663(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=99356017; PubMed=10426958; DOI=10.1038/11563;
RA Aevermann A., Seger K., Turley S., Sokatch J.R., Hol W.G.J.;
RT "Crystal structure of 2-oxoisovalerate and dehydrogenase and the
RT architecture of 2-oxo acid dehydrogenase multienzyme complexes.";
RL Nat. Struct. Biol. 6:785-792(1999).
CC -1- FUNCTION: The branched-chain alpha-keto dehydrogenase complex
CC catalyzes the overall conversion of alpha-keto acids to acyl-CoA
CC and CO(2). It contains multiple copies of three enzymatic
CC components: branched-chain alpha-keto acid decarboxylase (E1),
CC lipamide acyltransferase (E2) and lipamide dehydrogenase (E3).
CC -1- CATALYTIC ACTIVITY: 3-methyl-2-oxobutanoate +
CC [dihydrolipoyl]lysine-residue (2-methylpropanoyl)transferase]
CC lipoyllysine = [dihydrolipoyl]lysine-residue (2-
CC methylpropanoyl)transferase] S-(2-
CC methylpropanoyl)dihydrolipoyllysine + CO(2).
CC -1- Cofactor: Thiamine pyrophosphate.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57613; AAA65614.1; -.
DR PIR; S01317; DEPSXA.
DR PDB; 1Q50; X-ray; A=2-408.
DR InterPro; IPR001017; Dehydrogenase_E1.
DR Pfam; PF00676; E1_dh; 1.
DR 3D-structure; Flavoprotein; Oxidoreductase; Thiamine pyrophosphate.
KW TURN 18
FT TURN 19
FT HELIX 24
FT TURN 26
FT TURN 32
FT TURN 33
FT TURN 40
FT HELIX 44
FT HELIX 47
FT HELIX 48
FT TURN 51
FT TURN 52
FT STRAND 55
FT STRAND 58
FT STRAND 60
FT TURN 61
FT STRAND 64
FT STRAND 67
FT HELIX 74
FT HELIX 99
FT TURN 100
FT TURN 101
FT TURN 110
FT HELIX 112
FT HELIX 113
FT TURN 125
FT STRAND 126
FT STRAND 128
FT STRAND 130
FT TURN 136
FT HELIX 141
FT TURN 142
FT HELIX 143
FT HELIX 146
FT TURN 155

```



```

FT TURN 157 158
FT TURN 160 163
FT TURN 167 168
FT STRAND 172 172
FT HELIX 173 175
FT TURN 176 176
FT STRAND 177 177
FT HELIX 186 200
FT TURN 201 202
FT STRAND 207 212
FT TURN 213 213
FT HELIX 214 217
FT TURN 219 231
FT STRAND 232 232
FT STRAND 235 241
FT STRAND 244 245
FT TURN 246 247
FT STRAND 248 249
FT HELIX 250 253
FT TURN 254 257
FT TURN 261 261
FT HELIX 262 266
FT TURN 267 268
FT STRAND 270 275
FT TURN 276 277
FT HELIX 279 294
FT TURN 295 296
FT STRAND 300 305
FT TURN 314 315
FT HELIX 318 320
FT TURN 321 321
FT TURN 324 325
FT HELIX 326 329
FT TURN 331 332
FT HELIX 335 345
FT TURN 346 347
FT HELIX 351 373
FT TURN 374 375
FT HELIX 388 390
FT HELIX 399 406
FT TURN 407 407
SQ SEQUENCE 410 AA; 45268 MW; 0C998460CCFB9CF4 CRC64;

Query Match 49.5%; Score 48.5; DB 1; Length 410;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 4 GPTLRWLTSTRT-PHS 18
DB 298 GPTLRWLTSTRT-PHS 313

RESULT 10
Q88EQ2 PRELIMINARY; PRT; 410 AA.
AC Q88EQ2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 2-oxoisovalerate dehydrogenase, alpha subunit.
GN Name=bkd1; OrderedLocNames=PP4401;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weiner C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfele B.K., Scanlan D., Tran K.,

```

```

RA Moszzer A., Uteerback T.R., Rizzo M., Lee K., Kosack D., Moesti D.,
RA Medler H., Lauber J., Sejpandic D., Hohelsel J., Straetz M., Heim S.,
RA Klewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Frazer C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016790; AAN69979.1; -.
DR HSSP; P09060; I080.
DR TIGR; PP4401; -.
DR GO; GO:0016624; F:oxidoreductase activity, acting on the alde. .; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001017; Dehydrogenase_E1.
DR Pfam; PF00676; El_dh; 1.
KW Complete proteome.
SQ SEQUENCE 410 AA; 45220 MW; B1AA98211D94A212 CRC64;

Query Match 49.5%; Score 48.5; DB 2; Length 410;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 4 GPTLRWLTSTRT-PHS 18
DB 298 GPTLRWLTSTRT-PHS 313

RESULT 11
Q91IM2 PRELIMINARY; PRT; 410 AA.
AC Q91IM2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 2-oxoisovalerate dehydrogenase (Alpha subunit).
GN Name=bkd1; OrderedLocNames=PA2247;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Golty L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laibig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL; AE004650; AAG05635.1; -.
DR PIR; C83365; C83365.
DR PIR; S05057; S05057.
DR HSSP; P09060; I080.
DR GO; GO:0016624; F:oxidoreductase activity, acting on the alde. .; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001017; Dehydrogenase_E1.
DR Pfam; PF00676; El_dh; 1.
KW Complete proteome.
SQ SEQUENCE 410 AA; 45256 MW; BE3AF6FAB6F0F01 CRC64;

Query Match 49.5%; Score 48.5; DB 2; Length 410;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 4 GPTLRWLTSTRT-PHS 18
DB 298 GPTLRWLTSTRT-PHS 313

RESULT 12

```

09RV54 PRELIMINARY; PRT; 150 AA.  
 ID 09RV54;  
 AC 09RV54;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Hypothetical protein DR0948.  
 GN OrderedLocustNames=DR0948;  
 OS Deinococcus radiodurans.  
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;  
 RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,  
 RA Fraser C.M.,  
 RA "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans R1.";  
 RT Science 286:1571-1577 (1999).  
 DR EMBL; AF001947; AAF10530.1; -.  
 DR PIR; C75456; C75456.  
 DR TIGR; DR0948; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 150 AA; 16891 MW; 69A7695F09BBF3FB CRC64;

Query Match 49.0%; Score 48; DB 2; Length 150;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Qy 2 IEGLPTLRWLTSTRTP 17  
 Db 97 LDGSPAREWQTEGTPH 112

RESULT 13  
 066272 PRELIMINARY; PRT; 245 AA.  
 ID 066272;  
 AC 066272;  
 DT 01-AUG-1998 (TReMBLrel. 07, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Photosynthetic reaction center L subunit (Fragment).  
 GN Name=pufl;  
 OS Erythrobacter litoralis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
 OC Sphingomonadaceae; Erythrobacter.  
 OX NCBI\_TaxID=39960;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAM4332;  
 RX MEDLINE=21622632; PubMed=11832943; DOI=10.1038/415630a;  
 RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,  
 RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.,  
 RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs";  
 RL Nature 415:630-633 (2002).  
 DR EMBL; AB010981; BAA25791.1; -.  
 DR HSSP; P02954; IQOV.  
 DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.  
 DR GO; GO:0045156; P:electron transporter, transferring electron. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.  
 DR InterPro; IPR005871; Photo\_L.  
 DR InterPro; IPR000484; Photo\_RC.  
 DR Pfam; PF00124; Photo\_RC; 1.  
 DR PRINTS; PRO0256; REACTNCENTRE.  
 DR TIGRFAMs; TIGR01157; pufl; 1.  
 DR TIGRFAMs; TIGR01157; pufl; 1.

DR PROSITE; PS00244; REACTION\_CENTER; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 245 AA; 27214 MW; 52B266713E199ABD CRC64;  
 Qy Query Match 49.0%; Score 48; DB 2; Length 245;  
 Best Local Similarity 56.2%; Pred. No. 20;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 SIEGLPTLRWLTSTRTP 16  
 25 AIEGLTPNFWLIDIQP 40

RESULT 14  
 082989 PRELIMINARY; PRT; 249 AA.  
 ID 082989;  
 AC 082989;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Photosynthetic reaction center L subunit (Fragment).  
 GN Name=pufl;  
 OS Erythrobacter sp.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
 OC Sphingomonadaceae; Erythrobacter.  
 OX NCBI\_TaxID=1042;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MBIC3019;  
 RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;  
 RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,  
 RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.,  
 RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs";  
 RL Nature 415:630-633 (2002).  
 DR EMBL; AB015708; BAA32995.1; -.  
 DR HSSP; P02954; IYST.  
 DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.  
 DR GO; GO:0045156; P:electron transporter, transferring electron. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.  
 DR InterPro; IPR005871; Photo\_L.  
 DR InterPro; IPR000484; Photo\_RC.  
 DR Pfam; PF00124; Photo\_RC; 1.  
 DR PRINTS; PRO0256; REACTNCENTRE.  
 DR TIGRFAMs; TIGR01157; pufl; 1.  
 DR PROSITE; PS00244; REACTION\_CENTER; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 249 AA; 27702 MW; 4D68EDC82B7166AD CRC64;

Query Match 49.0%; Score 48; DB 2; Length 249;  
 Best Local Similarity 56.2%; Pred. No. 20;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 SIEGLPTLRWLTSTRTP 16  
 25 AIEGLTPNFWLIDIQP 40

RESULT 15  
 09XDVO PRELIMINARY; PRT; 278 AA.  
 ID 09XDVO;  
 AC 09XDVO;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Photosynthetic reaction center L subunit.  
 GN Name=pufl;  
 OS Erythrobacter sp. MBIC3960.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
 OC Sphingomonadaceae; Erythrobacter.  
 OX NCBI\_TaxID=94771;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=MBIC3960;  
 RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;  
 RA Befa O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,  
 RA Hamada T., Eisen J.A., Frazer C.M., DeLong E.F.;  
 RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs";  
 RL Nature 415:630-633 (2002).  
 DR EMBL; AB027515; BAA78672.1; -;  
 DR HSSP; P02954; 1YST.  
 DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.  
 DR GO; GO:0045156; P:electron transporter, transferring electron. . .; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.  
 DR InterPro: IPR005871; Photo\_L.  
 DR InterPro: IPR000484; Photo\_RC.  
 DR Pfam; PF00124; Photo\_RC; 1.  
 DR PRINTS; PRO0256; REACTCENTRE.  
 DR TIGRPFAM; TIGR01157; pufl; 1.  
 DR PROSITE; PS00244; REACTION CENTER; 1.  
 SQ SEQUENCE 278 AA; 30735 MW; 0BE618844B3C54FB CRC64;

Query Match 49.0%; Score 48; DB 2; Length 278;  
 Best Local Similarity 56.2%; Pred. No. 23;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SEGPLREWLTSRTP 16  
 :|||||  
 DB 54 AIEGPTLNPMLIDIQ 69

## RESULT 16

Q9RKM5 PRELIMINARY; PRT; 319 AA.

AC Q9RKM5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative MerR family transcriptional regulator.  
 GN ORFNames=SCD17.06c;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NCBI\_TaxId=1902;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Frazer A., Godle A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Lake L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinovitch E., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)";  
 RL Nature 417:141-147 (2002).  
 CC -1- SIMILARITY: Contains 1 HTH merR-type DNA-binding domain.  
 DR EMBL; AL939118; CAB56383.1; -;  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR000551; HTH\_MerR.  
 DR InterPro: IPR009061; Putativ\_DNA\_bind.  
 DR Pfam; PF00376; MerR; 1.  
 DR PRINTS; PRO0040; HTHMER.  
 DR SMART; SM00422; HTH\_MER\_1.  
 DR PROSITE; PS50937; HTH\_MER\_2; 1.  
 SQ Complete proteome; DNA-binding.  
 OS SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365B CRC64;

Query Match 49.0%; Score 48; DB 2; Length 319;

Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 SEGPLREWLTSR 14  
 :|||||  
 DB 258 DSEPLREWLGR 269

## RESULT 17

Q9KX10 PRELIMINARY; PRT; 388 AA.

AC Q9KX10;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Protein kinase.  
 GN Name=pKnb; Synonyms=ORF388;  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 NCBI\_TaxId=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COL;  
 RX MEDLINE=20031141; PubMed=10566865;  
 RA de Lencastre H., Wu S.W., Pinho M.G., Lindovics A.M., Filipe S.,  
 RA Gardete S., Sobral R., Gill S., Chung M., Tomasz A.;  
 RT "Antibiotic resistance as a stress response: complete sequencing of a  
 large number of chromosomal loci in Staphylococcus aureus strain COL  
 that impact on the expression of resistance to methicillin";  
 RL Microb. Drug Resist. 5:163-175 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COL;  
 RA Wu S.;

Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; Y13639; CAA73979.1; -;  
 DR HSSP; P71584; 106Y.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_Thr\_kinase.  
 DR InterPro: IPR008271; Ser\_Thr\_pkin\_AS.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 388 AA; 43764 MW; 0582809E06379580 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 388;  
 Best Local Similarity 58.8%; Pred. No. 33;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTPS 18  
 :|||||  
 DB 90 IEGPLSYEYIESHGPLS 106

## RESULT 18

Q8GA10 PRELIMINARY; PRT; 481 AA.

AC Q8GA10;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative amino acid permease.  
 OS Arthrobacter nicotovorans.  
 OG Plasmid pAO1.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Micrococciaceae; Micrococcaceae; Arthrobacter.  
 OK NCBI\_TaxID=29320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95115562; PubMed=7815950;  
 RA Grether-Beck S., Igloi G.L., Pust S., Schletz R., Decker K.,  
 RA Brandesch R.;  
 RA "Structural analysis and molybdenum-dependent expression of the PAO1-  
 RT encoded nicotine dehydrogenase genes of Arthrobacter nicotinovorans.",  
 RL Mol. Microbiol. 13:929-936(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96172783; PubMed=8588735;  
 RA Menendez C., Igloi G., Hemminger H., Brandesch R.;  
 RT "A PAO1-encoded molybdopterin cofactor gene (moa) of Arthrobacter  
 RT nicotinovorans: characterization and site-directed mutagenesis of the  
 RT encoded protein.",  
 RL Arch. Microbiol. 164:142-151(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98088982; PubMed=9428706;  
 RA Menendez C., Otto A., Igloi G., Nick P., Brandesch R., Schubach B.,  
 RA Botcher B., Brandesch R.;  
 RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a  
 RT bacterial plasmid. Characterization of MoaA as a filament-forming  
 RT protein with adenosinetriphosphatase activity.",  
 RL Eur. J. Biochem. 250:524-531(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99096870; PubMed=9878353; DOI=10.1006/jmbi.1998.2227;  
 RA Schenk S., Hoelz A., Kraus B., Decker K.;  
 RT "Gene structure and properties of enzymes of the plasmid-encoded  
 RT nicotine catabolism of Arthrobacter nicotinovorans.",  
 RL J. Mol. Biol. 284:1323-1339(1998).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97230479; PubMed=9073580; DOI=10.1006/plas.1996.1272;  
 RA Menendez C., Igloi G.L., Brandesch R.;  
 RT "IS1473, a putative insertion sequence identified in the plasmid PAO1  
 RT from Arthrobacter nicotinovorans: isolation, characterisation and  
 RT distribution among Arthrobacter species.",  
 RL Plasmid 37:35-41(1997).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21405725; PubMed=11514508;  
 RA DOI=10.1128/JB.183.18.5262-5267.2001;  
 RA Baisach D., Sandu C., Brandesch R., Igloi G.L.;  
 RT "A gene cluster on pAO1 of Arthrobacter nicotinovorans involved in the  
 RT degradation of the plant alkaloid nicotine: cloning, purification and  
 RT characterization of 2,6-dihydroxytryptidine 3-hydroxylase.",  
 RL J. Bacteriol. 183:5262-5267(2001).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22505657; PubMed=12618462;  
 RA DOI=10.1128/JB.185.6.1976-1986.2003;  
 RA Igloi G.L., Brandesch R.;  
 RT "Sequence of the 16S-kilobase catabolic plasmid PAO1 from Arthrobacter  
 RT nicotinovorans and identification of a PAO1-dependent nicotine uptake  
 RT system.",  
 RL J. Bacteriol. 185:1976-1986(2003).  
 RN [8]  
 RP EMBL; AJ507836; CAD47924.1; -;  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005279; F: amino acid-polyamine transporter activity; IEA.  
 DR GO; GO:0006865; P: amino acid transport; IEA.  
 DR GO; GO:0006810; P: transport; IEA.  
 DR InterPro; IPR0062293; AA/perl\_permease1.  
 DR InterPro; IPR004841; Permease\_region.  
 DR Pfam; PF00324; AA\_permease; I.  
 KM Plasmid; Transmembrane; Transport.  
 SO SEQUENCE 481 AA; 49782 MW; 4EA9FB3BB8B76B64 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 481;

Best Local Similarity 56.2%; Pred. No. 42;  
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 SIEGPTLRWLTSTRP 16  
 Db 366 SLVGPVWMLWLSSTP 381  
 RESULT 19  
 ID 0751J9 PRELIMINARY; PRT; 632 AA.  
 AC 0751J9;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein B1130610.3 (Hypothetical protein  
 DE P0022D06.17).  
 GN Name=B1130610.3; Synonym=P0022D06.17;  
 OS Oryza sativa [Japonica cultivar-group].  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrharioideae; Oryzoae; Oryza.  
 OK NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Lee P.-F., Chang S.-J., Chen H.-C., Chen S.-K.,  
 RA Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y.,  
 RA Hsiao S.-H., Hsiung J.-N., Hsu C.-H., Kuo P.-I., Lee M.-C.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.;  
 RT Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Chang S.-J., Chung C.-H., Chung C.-I., Han S.-Y.,  
 RA Hsiao S.-H., Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.;  
 RT "Oryza sativa PAC P0022D06 genomic sequence.",  
 RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AC130603; AAT01306.1; -;  
 DR EMBL; AC132485; AAU03115.1; -;  
 DR InterPro; IPR009105; Colicin\_E3\_cat.  
 KW Hypothetical protein.  
 SO SEQUENCE 632 AA; 69035 MW; 8EBDE6377EDB5D0F CRC64;  
 Query Match 49.0%; Score 48; DB 2; Length 632;  
 Best Local Similarity 60.0%; Pred. No. 56;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 IEGPTLRWLTSTRP 16  
 Db 400 LEGESLREWLFPDTP 414  
 RESULT 20  
 ID 08NX14 PRELIMINARY; PRT; 664 AA.  
 AC 08NX14;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE MW1103 protein.  
 GN OrderedLocustNames=MW1103;  
 OS Staphylococcus aureus (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OK NCBI\_TaxID=196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MW2;  
 RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;

```

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RU Lancet 359:1819-1827(2002).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL, AF004826; BAB94968.1; -.
DR HSP, P71584; 106Y.
DR GO: GO:0005524; P:ATP binding; IEA.
DR GO: GO:0008658; P:penicillin binding; IEA.
DR GO: GO:0006674; P:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR001009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF03793; PASTA_2.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00740; PASTA_3.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Complete proteome; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 664 AA; 74363 MW; 26F1386C5DB61828 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 664;
Best Local Similarity 58.8%; Pred. No. 59;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTPS 18
DB 90 IEGPTLRWLTSTRTPS 106
ID ||||| |::| |
AC 090P8 PRELIMINARY; PRT; 664 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein kinase.
GN OrderedLocusNames=SAV1220;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsunari H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shida T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RU Lancet 357:1225-1240(2001).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL, AF003361; BAB57382.1; -.
DR PIR, G89894; G89894.
DR HSP, P71584; 106Y.
DR GO: GO:0005524; P:ATP binding; IEA.
DR GO: GO:0006674; P:protein serine/threonine kinase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.

```

```

DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR005543; PASTA_1.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF03793; PASTA_2.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00740; PASTA_3.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Complete proteome; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 664 AA; 74377 MW; 3461386C5DB61828 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 664;
Best Local Similarity 58.8%; Pred. No. 59;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPTLRWLTSTRTPS 18
DB 90 IEGPTLRWLTSTRTPS 106
ID ||||| |::| |
AC 07A528 PRELIMINARY; PRT; 664 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Protein kinase.
GN OrderedLocusNames=SA1063;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsunari H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shida T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RU Lancet 357:1225-1240(2001).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL, AF003361; BAB42315.1; -.
DR GO: GO:0005524; P:ATP binding; IEA.
DR GO: GO:0008658; P:penicillin binding; IEA.
DR GO: GO:0006674; P:protein serine/threonine kinase activity; IEA.
DR GO: GO:0006468; P:transferase activity; IEA.
DR GO: GO:0016740; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR005543; PASTA_1.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF03793; PASTA_2.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00740; PASTA_3.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

```

KW ATP-binding: Complete proteome; Kinase;  
 KM Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 664 AA; 74377 MW; 3461386C5DB61828 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 664;  
 Best Local Similarity 58.8%; Pred. No. 59;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTPHS 18  
 ||||| |::| |  
 Db 90 IEGPLSEYIESHGPLS 106

RESULT 23

Q6G9Z3 PRELIMINARY; PRT; 664 AA.

AC Q6G9Z3; SEQUENCE FROM N.A.  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Serine/threonine-protein kinase (EC 2.7.1.1-).  
 GN OrderedLocustNames=SA51154;  
 OS Staphylococcus aureus (strain MSSA476).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxId=282459;

RA Spratt B.G., Parkhill J.;  
 RA "Complete genomes of two clinical Staphylococcus aureus strains:  
 RT evidence for the rapid evolution of virulence and drug resistance.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; BX571857; CAG42931.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0006658; F:penicillin binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR005543; PASA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_kin\_AS.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam; PF03793; PASTA; 2.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00740; PASTA; 3.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding: Complete proteome; Kinase;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 664 AA; 74363 MW; 26F1386C5DB61828 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 664;  
 Best Local Similarity 58.8%; Pred. No. 59;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTPHS 18  
 ||||| |::| |

Db 90 IEGPLSEYIESHGPLS 106

RESULT 24

Q6GHL5 PRELIMINARY; PRT; 664 AA.

AC Q6GHL5; SEQUENCE FROM N.A.  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Serine/threonine-protein kinase (EC 2.7.1.1-).  
 GN Name=pxnb; OrderedLocustNames=SA81196;  
 OS Staphylococcus aureus (strain MRSA252).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxId=282459;

RA Spratt B.G., Parkhill J.;  
 RA "Complete genomes of two clinical Staphylococcus aureus strains:  
 RT evidence for the rapid evolution of virulence and drug resistance.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; BX571856; CAG40198.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0006658; F:penicillin binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR005543; PASA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_kin\_AS.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam; PF03793; PASTA; 2.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00740; PASTA; 3.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding: Complete proteome; Kinase;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 664 AA; 74363 MW; 26F1386C5DB61828 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 664;  
 Best Local Similarity 58.8%; Pred. No. 59;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IEGPLREWLTSRTPHS 18  
 ||||| |::| |  
 Db 90 IEGPLSEYIESHGPLS 106

RESULT 25

Q8Y015 PRELIMINARY; PRT; 91 AA.

AC Q8Y015; SEQUENCE FROM N.A.  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein Rsc1059.

```

GN Name=RS0149; OrderedLocNames=RS01059;
OS Ralstonia solanacearum (pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCB1_TaxId=305;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000; PubMed=11823852; DOI=10.1038/415497a;
RX MEDLINE=21681879;
RA Salanoubat M., Genin S., Artiguenave F., Gouy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AF646062; CAD14761.1; -.
KW Complete proteome.
SQ SEQUENCE 91 AA; 10321 MW; 2B4DFEB37A528AD CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 2; Length 91;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGGPTLRWLTSTRTP 16
Db 75 LDGPAVQAWMLAQTP 89

RESULT 26
YD55_HABIN STANDARD; PRT; 154 AA.
AC P4168;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical UPF0260 protein H11355.
GN OrderedLocNames=H11355;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCB1_TaxId=727;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shrivley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudel D.M., Brandon R.C.,
RA Pine L.D., Frichman J.A., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512 (1995).
CC -1- SIMILARITY: Belongs to the UPF0260 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32814; AAC23002.1; -.
CC PIR; P64026; P64026.
CC TIGR; H11355; -.
CC HAMAP; MF_00676; -. 1.

```

```

DR InterPro; IPR008228; UCP006173.
DR Pfam; PF05779; DUF838; 1.
DR PIRSF; PIRSF006173; UCP006173; 1.
DR PRODOM; PD021710; UCP006173; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 154 AA; 18163 MW; 886CE6D467E8AB55 CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 1; Length 154;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 3 EGGPTLRWLTSTRTPHS 18
Db 105 EGGPTLRWHLTSGSPHS 122

RESULT 27
O66278 PRELIMINARY; PRT; 245 AA.
ID O66278;
AC O66278;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=psufL;
OS Agrobacterium sanguineum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae.
OX NCB1_TaxId=73269;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JAM12620;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633 (2002).
DR EMBL; AB011074; BAA25722.1; -.
DR HSSP; P02954; 1QOV;
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . . IEA.
DR GO; GO:0045156; P:electron transporter, transferring electron. . . IEA.
DR GO; GO:006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTCENTRE.
DR TIGRPFAM; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 245 AA; 26840 MW; DBACDB4DA050DB80 CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 2; Length 245;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 EGGPTLRWLTSTRTPHS 18
Db 27 QGGPTLRWHLTSGSPHS 42

RESULT 28
Q70BP7 PRELIMINARY; PRT; 282 AA.
ID Q70BP7;
AC Q70BP7;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE AgCP2340 (Fragment).
GN Name=agCG44337; ORFNames=ENSANGG00000014770;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematoceera; Culicoides; Anopheles.

```



```

OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA0100878; EAA08303.1; -.
DR InterPro; IPR008042; Retrotrans_Pao.
DR Pfam; PF05380; Peptidase_A17; 1.
FT NON_TER 1
FT NON_TER 282
SQ SEQUENCE 282 AA; 31558 MW; 8154BA517F1FD32A CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 2; Length 282;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 STEGPTLRKWLTSRT 15
DB 31 SLLEGQLQEWLQFRT 45

RESULT 29
Q66CD5 PRELIMINARY; PRT; 417 AA.
AC Q6J6D5;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Carboxypeptidase A (EC 3.4.17.1).
OS Name=CpA-VI;
GN Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RA Iseo J., Amenezes A., Wells M.A.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY590492; AAT36730.1; -.
DR HSSP; P00730; IABM.
DR GO; GO:0004182; F:carboxypeptidase A activity; IEA.
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000834; Peptidase_M14.
DR InterPro; IPR003146; Prot_inh_M14.
DR InterPro; IPR009020; Prot_inh_propept.
DR Pfam; PF00246; Peptidase_M14; 1.
DR Pfam; PF02244; Propep_M14; 1.
DR PRINTS; PR00765; CRBOXYPTRASEA.
DR SMART; SM00631; Zn_pept; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR Carboxypeptidase; Hydrolase.
SQ SEQUENCE 417 AA; 47505 MW; E599DB6D4A1B97B CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 2; Length 417;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 TLREWLTSRTPHS 18
DB 226 TNRQWRKTRTPHS 228

RESULT 30
Q31703 PRELIMINARY; PRT; 430 AA.
AC Q31703; Q7BVS3;

```

```

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Molybdopterin biosynthesis protein MoeA.
GN Name=moeA; OrderedLocustNames=BSU14280;
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=9604033; PubMed=9384377; DOI=10.1038/36786;
RA Kunat F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bortero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
RA Broillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Deutz F., Devine K.M., Distelhof A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Hénaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Kleber-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara K., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Potwolik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivoita C., Rocha E., Roche B.,
RA Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekoska A., Serró S.J., Serró P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takaki T., Takahashi H.,
RA Takemaru K., Takeuchi M., Yamakoshi A., Yanaka I., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandendol M., Vanlier F.,
RA Vassarotti A., Viari A., Wambut R., Wedler E., Wedler F.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1168;
RX MEDLINE=90368558; PubMed=1697575;
RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.,
RT "Secretory S complex of Bacillus subtilis: sequence analysis and
RT identity to pyruvate dehydrogenase."
RL J. Bacteriol. 172:5052-5063(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1168;
RX MEDLINE=97144523; PubMed=8990290;
RA Heniques A.O., Bryan E.M., Beall B.W., Moran C.P., Jr.;
RT "case1, case60, and case22 are new members of mother-cell-specific
RT sporulation regulons in Bacillus subtilis."
RL J. Bacteriol. 179:389-398(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=1168;
RA Calwell R.M., Ferrari E.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99111; CAB13301.1; -.
DR EMBL; AF012285; AAC24902.1; -.
DR PIR; B69659; B69659.
DR HSSP; P12281; 1G8L.
DR GO; GO:0006777; P:Mo-molybdopterin cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MoCF_biosynth.
DR InterPro; IPR005111; MoEA_C.
DR InterPro; IPR005110; MoEA_N.

```



DR Pfam; PRO0994; MOCF\_biosynth; 1.  
 DR Pfam; PF03454; MoeA\_C; 1.  
 DR Pfam; PF03453; MoeA\_N; 1.  
 DR ProDom; PD002460; MOCF\_biosynth; 1.  
 DR TIGRFAMs; TIGR00177; molyb\_syn; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 430 AA; 46619 MW; DBE2C5FE9F542388 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 430;  
 Best Local Similarity 50.0%; Pred. No. 53;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 PTLREWLTSRTPHS 18  
 Db 323 PTLQWLNLNTRPHS 336

RESULT 31  
 O8CSV9 PRELIMINARY; PRT; 667 AA.  
 AC O8CSV9;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Protein kinase.  
 GN OrderedLocustNames=SE0895;  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 NX NCBI\_TaxID=1282;  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RX PubMed=12950922;  
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
 RA Qiu Z.-Q., Zhao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
 RA Yuan Z.-H., Zhao G.-P., Gu D., Danchin A., Wen Y.-M.;  
 RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
 RT Staphylococcus epidermidis strain (ATCC 12228).";  
 RL Mol. Microbiol. 49:1577-1593(2003).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AE016746; AAC04492.1; -.  
 DR HSSP; P71584; 106Y  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008658; F:penicillin binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011008; Kinase\_like.  
 DR InterPro; IPR005543; PASTA.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam; PF03793; PASTA; 2  
 DR Pfam; PF00069; Kinase; 1  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00740; PASTA; 3.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Complete proteome; Kinase;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 667 AA; 75411 MW; 479877B4531CD97 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 667;  
 Best Local Similarity 58.8%; Pred. No. 86;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IEQPTLREWLTSRTPHS 18  
 Db 90 IEQPTLAETIHSHPHS 106

## RESULT 32

ID Q7MWY0 PRELIMINARY; PRT; 818 AA.  
 AC Q7MWY0;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative TrbE-like protein.  
 GN Name=trbE; ORFNames=PHG362;  
 OS Alkaliigenes eutrophus (Ralstonia eutropha).  
 OG Alkaligenes eutrophus subsp. PHG1.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Wautersia.  
 NX NCBI\_TaxID=510;  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=H16;  
 RX MEDLINE=22830147; PubMed=12948488; DOI=10.1016/S0022-2836(03)00894-5;  
 RA Schwartz E., Henne A., Cramm R., Bittinger T., Friedrich B.,  
 RA Gotschalk G.;  
 RT "Complete Nucleotide Sequence of PHG1: A Ralstonia eutropha H16  
 RT Megaplasmid Encoding Key Enzymes of H2-based Lithoautotrophy and  
 RT Anaerobiosis.";  
 RL J. Mol. Biol. 332:369-383(2003).  
 DR EMBL; AY305378; AAP8611.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR InterPro; IPR004346; CAGE\_TrbE\_VtrB.  
 DR Pfam; PF03135; CAGE\_TrbE\_VtrB; 1.  
 KW Plasmid.  
 SQ SEQUENCE 818 AA; 92392 MW; AFD46E761EDC99B CRC64;

Query Match 48.0%; Score 47; DB 2; Length 818;  
 Best Local Similarity 50.0%; Pred. No. 11e+02;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IEQPTLREWLTSRT 15  
 Db 426 IEQPTLREWLTSRT 439

## RESULT 33

ID ODBA\_BACSU STANDARD; PRT; 330 AA.  
 AC P37940;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE 2-oxoisovalerate dehydrogenase alpha subunit (EC 1.2.4.4) (Branched-  
 DE chain alpha-keto acid dehydrogenase E1 component alpha chain) (BCKDH  
 DE E1-alpha).  
 GN Name=bfmbA; Synonyms=bfmbA; OrderedLocustNames=BSU24050;  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 NX NCBI\_TaxID=1423;  
 RX SEQUENCE FROM N.A., AND SEQUENCE OF 1-22.  
 RC STRAIN=168;  
 RX MEDLINE=93279308; PubMed=8504804;  
 RA Wang G.-F., Kuriki T., Roy K.L., Kaneda T.;  
 RT "The primary structure of branched-chain alpha-oxo acid dehydrogenase  
 RT from Bacillus subtilis and its similarity to other alpha-oxo acid  
 RT dehydrogenases.";  
 RL Eur. J. Biochem. 213:1091-1099(1993).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=97124195; PubMed=8969508;  
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
 RA Kobayashi Y.;  
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
 RT the Bacillus subtilis genome containing the skin element and many  
 RT sporulation genes.";  
 RL Microbiology 142:3103-3111(1996).

[3]  
SEQUENCE FROM N.A.  
RX STRAIN=168;  
RX MEDLINE=98044033; PubMed=93843577; DOI=10.1038/36786;  
RA Kuner F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,  
RA Brouillet S., Brunsch C.V., Caldwell B., Campano V., Carter N.M.,  
RA Choi S.K., Codart J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devane K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.-D., Errington J., Fabrit C., Ferrari E., Foubier D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,  
RA Ghim S.-Y., Glaser P., Goffeau A., Goltigly E.J., Grandi G.,  
RA Giuseppe G., Guy B.-J., Haga K., Haeleth J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holteppel S., Hosono S., Hullo M.F., Itaya M.,  
RA Jones L.-M., Joris B., Karamata P., Kasahara T., Kleier-Blanchard M.,  
RA Klein C., Kobayashi Y., Koelter P., Koningsstein G., Krogh S.,  
RA Kunano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
RA Lazarevic V., Lee S.M., Levine A., Liu H., Maeda S., Manuel C.,  
RA Medigne C., Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S.,  
RA Nodack M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B.,  
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Potwollik S.,  
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,  
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rooba E., Roche B.,  
RA Rose M., Sadie Y., Sato T., Scanlan A., Sero S.J., Serron P.,  
RA Scoffone F., Sekiguchi J., Sekowska A., Sero S.J., Serron P.,  
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
RA Takemaru K., Takeuchi M., Yamakoshi A., Tanaka T., Terpetra P.,  
RA Tognoni A., Tosato V., Uchiyama S., Vandenoil M., Vanlier F.,  
RA Vassarotti A., Viari A., Wambuit R., Wedler E., Wedler H.,  
RA Wetzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,  
RA Yoshikawa H., Danchin A.,  
RT "the complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"  
RT Nature 390:249-256(1997).  
CC -1- FUNCTION: The branched-chain alpha-keto dehydrogenase complex  
CC catalyzes the overall conversion of alpha-keto acids to acyl-CoA  
CC and CO(2). It contains multiple copies of three enzymatic  
CC components: branched-chain alpha-keto acid decarboxylase (E1),  
CC lipamide acyltransferase (E2) and lipamide dehydrogenase (E3).  
CC -1- CATALYTIC ACTIVITY: 3-methyl-2-oxobutanoate +  
CC [dihydrolipoyllysine-residue (2-methylpropionyl)transferase]  
CC 1lipoyllysine = [dihydrolipoyllysine-residue (2-  
CC methylpropionyl)transferase] S-(2-  
CC methylpropionyl)dihydrolipoyllysine + CO(2).  
CC -1- COFACTOR: Thiamine pyrophosphate.  
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, M97391, AAA22278.1, -,  
DR EMBL, D84432, BAA12598.1, -,  
DR EMBL, Z99116, CAA14336.1, -,  
DR PIR, C69593, C69593.  
DR HSSP, P12694, 1DTW.  
DR Subtilast, BG10307, bfmBA.  
DR InterPro, IPR001017, Dehydrogenase\_E1.  
DR Pfam, PF00676, E1\_dh, 1.  
KW Complete proteome; Direct protein sequencing; Flavoprotein;  
KW Oxidoreductase; Thiamine pyrophosphate.  
SQ SEQUENCE 330 AA; 36334 MW; 39584D3FA363B656 CRC64;

QY	3	EGPITLREWLTSR--TPHS 18	DB	237	EGPITLTIETISYRLTPHS 253
DB	237	EGPITLTIETISYRLTPHS 253			
RESULT 34					
ID	Q65HK8	PRELIMINARY;	PRT;	330 AA.	
AC	Q65HK8;				
DT	25-OCT-2004 (TREMBLrel. 28, Created)				
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)				
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)				
DE	BkdA (Branched-chain alpha-keto acid dehydrogenase E1 subunit) (2-oxoisovalerate dehydrogenase alpha subunit).				
GN	Name=bkdA; ORFNames=BL01504, BL02582;				
OS	Bacillus licheniformis DSM 13.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=279010;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=DSM 13;				
RX	PubMed=15383718;				
RA	Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H., Ehrenreich P., Baumeister S., Henne A., Liesegang H., Merkl R., Ehrenreich A., Gotschalk G.,				
RA	"The Complete Genome Sequence of Bacillus licheniformis DSM13, an Organism with Great Industrial Potential.";				
RL	J. Mol. Microbiol. Biotechnol. 7:204-211(2004).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 14580;				
RA	Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J., Tang M., de Leon A.L., Xiang H., Guest V., Clausen I.G., Olsen P.B., Raemussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D., Berkta R.M.;				
RA	"Complete genome sequence of the industrial bacterium Bacillus licheniformis and comparisons with closely related Bacillus species.";				
RL	Genome Biol. 5:R77-R77(2004).				
DR	EMBL; AE017333; AAU41456.1; -				
DR	EMBL; CP000002; AAU24056.1; -				
SQ	SEQUENCE 330 AA; 36463 MW; BR314979P9065C9B CRC64;				
QY	3	EGPITLREWLTSR--TPHS 18			
DB	237	EGPITLTIETISYRLTPHS 253			
RESULT 35					
ID	Q25564	PRELIMINARY;	PRT;	527 AA.	
AC	Q25564;				
DT	01-JAN-1998 (TREMBLrel. 05, Created)				
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Hypothetical protein HP0906.				
GN	OrderedLocusNames=HP0906;				
OS	Helicobacter pylori (Campylobacter pylori).				
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;				
OC	Helicobacteraceae; Helicobacter.				
OX	NCBI_TaxID=210;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=26695 / ATCC 700392;				
RX	MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;				
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S.R., Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,				

RA Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,  
 RA Hickey E.K., Berg D.B., Gocayne J.D., Utterback T.R., Peterson J.D.,  
 RA Kelley J.M., Cotton M.D., Maitman J.F., Fujii C., Bowman C.,  
 RA Wootley L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,  
 RA Smith H.O., Fraser C.M., Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."  
 RT Nature 388:539-547(1997).  
 RL EMBL; A5000600; AAD07958.1; --  
 DR PIR; B64633; B64633.  
 DR TIGR; H0906; --  
 DR CO; GO:0009424; C:flagellar hook (sensu Bacteria); IEA.  
 DR CO; GO:0003774; P:motor activity; IEA.  
 DR CO; GO:0009296; P:flagellum biogenesis; IEA.  
 DR InterPro; IPR001635; Flag\_hook.  
 DR Pfam; PF02120; Flag\_hook; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 527 AA; 58160 MW; C36FCB03D7FAE98 CRC64;

Query Match 47.4%; Score 46.5; DB 2; Length 156;  
 Best Local Similarity 44.4%; Pred. No. 80;  
 Matches 8; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 3 ESWTLEWLSR---TPH 17  
 DB 98 QAPTLKDWLNHKKTTTPH 115

## RESULT 36

064K11 PRELIMINARY; PRT; 156 AA.

AC 064K11;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE NADH dehydrogenase subunit II (Fragment).  
 GN Name=ND2;  
 OS Eleutherodactylus angustii (barking frog).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Leptodactylidae;  
 OC Telmatobiinae; Eleutherodactylus.  
 OK NCBI\_TaxID=228429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MF3813.TX;  
 RA Goldberg C.S., Sullivan B.K., Malone J.H., Schwalbe C.R.;  
 RT "Divergence among barking frogs (*Eleutherodactylus angustii*) in the  
 RT southwestern United States."  
 RL Herpetologica 60:312-320(2004).  
 DR EMBL; AY442954; AAS49146.1; --  
 KM Mitochondrion.  
 FT NON TER 1  
 SQ SEQUENCE 156 AA; 17094 MW; FB9CBB4507CF373A CRC64;

Query Match 46.9%; Score 46; DB 2; Length 156;  
 Best Local Similarity 80.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 ESWTSRTPHS 18  
 DB 77 EWLISSTPHS 86

## RESULT 37

064K12 PRELIMINARY; PRT; 156 AA.

AC 064K12;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE NADH dehydrogenase subunit II (Fragment).  
 GN Name=ND2;

OS Eleutherodactylus angustii (barking frog).

OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Leptodactylidae;  
 OC Telmatobiinae; Eleutherodactylus.  
 OK NCBI\_TaxID=228429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MF3811.TX, and MF3807.TX;  
 RA Goldberg C.S., Sullivan B.K., Malone J.H., Schwalbe C.R.;  
 RT "Divergence among barking frogs (*Eleutherodactylus angustii*) in the  
 RT southwestern United States."  
 RL Herpetologica 60:312-320(2004).  
 DR EMBL; AY442853; AAS49145.1; --  
 DR EMBL; AY442951; AAS49143.1; --  
 KM Mitochondrion.  
 FT NON TER 1  
 SQ SEQUENCE 156 AA; 17076 MW; 0B9CBB4507CF373B CRC64;

Query Match 46.9%; Score 46; DB 2; Length 156;  
 Best Local Similarity 80.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 ESWTSRTPHS 18  
 DB 77 EWLISSTPHS 86

## RESULT 38

064K16 PRELIMINARY; PRT; 156 AA.

AC 064K16;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE NADH dehydrogenase subunit II (Fragment).  
 GN Name=ND2;  
 OS Eleutherodactylus angustii (barking frog).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Leptodactylidae;  
 OC Telmatobiinae; Eleutherodactylus.  
 OK NCBI\_TaxID=228429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sonora.MX;  
 RA Goldberg C.S., Sullivan B.K., Malone J.H., Schwalbe C.R.;  
 RT "Divergence among barking frogs (*Eleutherodactylus angustii*) in the  
 RT southwestern United States."  
 RL Herpetologica 60:312-320(2004).  
 DR EMBL; AY442949; AAS49141.1; --  
 KM Mitochondrion.  
 FT NON TER 1  
 SQ SEQUENCE 156 AA; 16805 MW; D223D924A9290B9B CRC64;

Query Match 46.9%; Score 46; DB 2; Length 156;  
 Best Local Similarity 80.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 ESWTSRTPHS 18  
 DB 77 EWLISSTPHS 86

## RESULT 39

064K24 PRELIMINARY; PRT; 156 AA.

AC 064K24;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE NADH dehydrogenase subunit II (Fragment).  
 GN Name=ND2;

```

OS Eleutherodactylus augusti (barking frog).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Leptodactylidae;
OC Telmatobiinae; Eleutherodactylus.
OK NCBI_TaxId=228429;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Huachuacm65_AZ, Santaritam6_AZ, and Huachuacm617_AZ;
RA Goldberg C.S., Sullivan B.K., Malone J.H., Schwalbe C.R.;
RT "Divergence among barking frogs (Eleutherodactylus augusti) in the
RT southwestern United States."
RL Herpetologica 60:312-320(2004).
DR EMBL; AY442941; AAS49133.1; -.
DR EMBL; AY442948; AAS49140.1; -.
DR EMBL; AY442938; AAS49130.1; -.
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 156 AA; 16791 MW; DC21B329322A77EB CRC64;

Query Match 46.9%; Score 46; DB 2; Length 156;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 EMLTSRTPHS 18
DB 77 EMLISTPHS 86

RESULT 40
064K25 PRELIMINARY; PRT; 156 AA.
AC 064K25;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
DE NADH dehydrogenase subunit II (Fragment).
GN Name=ND2;
OS Eleutherodactylus augusti (barking frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Leptodactylidae;
OC Telmatobiinae; Eleutherodactylus.
OK NCBI_TaxId=228429;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Huachuacm630_AZ;
RA Goldberg C.S., Sullivan B.K., Malone J.H., Schwalbe C.R.;
RT "Divergence among barking frogs (Eleutherodactylus augusti) in the
RT southwestern United States."
RL Herpetologica 60:312-320(2004).
DR EMBL; AY442940; AAS49132.1; -.
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 156 AA; 16830 MW; DC21B329322A60DB CRC64;

Query Match 46.9%; Score 46; DB 2; Length 156;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 EMLTSRTPHS 18
DB 77 EMLISTPHS 86

RESULT 41
082PX5 PRELIMINARY; PRT; 377 AA.
AC 082PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
KW Hypothetical protein.

```

```

GN OrderedLocustNames=SAV747;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OK NCBI_TaxId=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211431198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005023; BAC68457.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 377 AA; 41307 MW; 0253176AAAB62F3 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 377;
Best Local Similarity 61.5%; Pred. No. 67;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IEPTLRLEWLSR 14
DB 168 MEGPDLRAWLPKR 180

RESULT 42
0745Z3 PRELIMINARY; PRT; 559 AA.
AC 0745Z3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
DE Ribonucleoside-diphosphate reductase alpha chain.
GN OrderedLocustNames=TRP0162;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OG Plaamid pT727.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OK NCBI_TaxId=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wietzer A., Hartach T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starckl V., Schlenker S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
RT thermophilus."
RL Nat. Biotechnol. 22:547-553(2004).
DR EMBL; AE017222; AAS82492.1; -.
DR GO; GO:0005971; Cytridonucleoside-diphosphate reductase complex; IEA.
DR GO; GO:0004748; P:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:006260; P:DNA replication; IEA.
DR InterPro; IPR000788; Ribonucleo_red.
DR InterPro; IPR010994; Ruva_2_like.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF02867; Ribonuc_red_1gc; 1.
DR PRINTS; PR01183; RIBORDTASEM1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Complete proteome.

```

SEQ SEQUENCE 559 AA; 63800 MW; E3980D88A831A6FB CRC64;

Query Match 46.9%; Score 46; DB 2; Length 559;  
Best Local Similarity 47.1%; Pred. No. 1e+02;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IEGLPTLEWLTSTRTPS 18  
Db 168 VERPDLLEWLSIOREHS 184

## RESULT 43

SYE\_PRRAS STANDARD; PRT; 570 AA.

ID SYE\_PRRAS STANDARD; PRT; 570 AA.  
AC Q82033;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
DE (Gluts).  
GN Name=glx; OrderedLocNames=PAE2969;  
OS Archaea: Crenarchaeota; Thermoprotei; Thermoproteales;  
OC Thermoproteaceae; Pyrobaculum.  
OX NCBI\_Taxid=13773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
RX MEDLINE=21664397; Pubmed=11792869; DOI=10.1073/pnas.241636498;  
RA Filiz-Gibdon S.T., Ladhner H., Kim U.-J., Stetter K.O., Simon M.I.,  
RA Miller J.H.;  
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
aerophilum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +  
CC diphosphate + L-glutamyl-tRNA(Glu).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase  
CC family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AE009904; AAL64575.1; -.  
DR HSSP; P00962; INYL.  
DR HAMAP; MF\_00022; -; 1.  
DR InterPro; IPR004526; Glx arch.  
DR InterPro; IPR000924; Glu tRNA-synt 1c.  
DR InterPro; IPR011035; Ribosomal\_L25rel.  
DR InterPro; IPR001412; tRNA-synt 1.  
DR Pfam; PF00749; tRNA-synt 1c; 1.  
DR Pfam; PF03950; tRNA-synt 1c; 1.  
DR PRINTS; PR00987; TRNASYNTHGLU.  
DR TIGRfam; TIGR00463; glx arch; 1.  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I, FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;  
KW Protein biosynthesis.  
FT SITE 107 117 "HIGH" region.  
SQ SEQUENCE 570 AA; 65837 MW; 767FCEB29A3064C CRC64;

Query Match 46.9%; Score 46; DB 1; Length 570;  
Best Local Similarity 42.1%; Pred. No. 1.1e+02;  
Matches 8; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

Qy 5 PTLREWL-----TSRTPH 17  
Db 258 PSYDWMVAFRITDTSKTPH 276

## RESULT 44

Q9K920 PRELIMINARY; PRT; 664 AA.

ID Q9K920 PRELIMINARY; PRT; 664 AA.  
AC Q9K920;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Serine/threonine protein kinase.  
GN OrderedLocNames=BH2504;  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_Taxid=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125;  
RX MEDLINE=20512582; Pubmed=11058132; DOI=10.1093/nar/28.21.4317;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
DR EMBL; AP001515; BAB06223.1; -.  
DR PIR; H83962; H83962.  
DR HSSP; P71584; 106Y.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008658; F:penicillin binding; IEA.  
DR GO; GO:0006574; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR011009; Kinase\_like.  
DR InterPro; IPR005543; PASTA.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF03793; PASTA; 3.  
DR Pfam; PF00069; PKinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00740; PASTA; 3.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Complete proteome; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 664 AA; 73719 MW; E2FP225DCC6BE52 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 664;  
Best Local Similarity 53.3%; Pred. No. 1.2e+02;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IEGLPTLEWLTSTRTP 16  
Db 90 VERPDLLEWLSIOREH 104

## RESULT 45

MYIC\_HUMAN STANDARD; PRT; 1028 AA.

ID MYIC\_HUMAN STANDARD; PRT; 1028 AA.  
AC O00159;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Myosin Ic (Myosin I beta) (MMI-beta) (MMIB).  
GN Name=MYOIC;  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 1, 2005, 15:48:02 ; Search time 87.3453 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-11

Perfect score: 106  
Sequence: 1 LAIBGPTLRQWLHGNGRDT 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_16Dec04:\*

1: \_geneseqp1980s:\*\n2: \_geneseqp1990s:\*\n3: \_geneseqp2000s:\*\n4: \_geneseqp2001s:\*\n5: \_geneseqp2002s:\*\n6: \_geneseqp2003as:\*\n7: \_geneseqp2003bs:\*\n8: \_geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	19	2	AAW09494
2	106	100.0	19	2	AAW09461
3	106	100.0	19	2	AAW36645
4	106	100.0	19	3	AAW17022
5	106	100.0	19	4	AAU25864
6	106	100.0	19	4	AAU25825
7	106	100.0	19	5	ABW72908
8	106	100.0	19	7	ADU73060
9	106	100.0	19	8	ADU52695
10	106	100.0	19	8	ADU51656
11	101	95.3	19	2	AAW33028
12	63	58.5	15	6	ABG71749
13	62	58.5	15	2	AAW67721
14	60.5	57.1	40	3	AAW17302
15	60	56.6	28	3	AAW17285
16	60	56.6	29	7	ADU73011
17	60	56.6	29	7	ADU73007
18	60	56.6	29	7	ADU73006
19	60	56.6	29	8	ADU52642
20	60	56.6	29	8	ADU52646
21	60	56.6	29	8	ADU52641
22	60	56.6	29	8	ADU51603
23	60	56.6	29	8	ADU51602
24	60	56.6	29	8	ADU51607
25	60	56.6	31	7	ADU73009

26	60	56.6	31	7	ADU73010	Adj73010 TPO mimet
27	60	56.6	31	8	ADU52644	Adj52644 CHI delet
28	60	56.6	31	8	ADU52645	Adj52645 CHI delet
29	60	56.6	31	8	ADU51606	Adj51606 CHI delet
30	60	56.6	31	8	ADU51605	Adj51605 CHI delet
31	59.5	56.1	18	7	ADU59659	Adn59659 Thrombopo
32	59.5	56.1	22	7	ADN59826	Adn59826 TWP pept1
33	59.5	56.1	25	7	ADN59700	Adn59700 Thrombopo
34	59.5	56.1	36	3	AAW96526	AAW96526 Thrombopo
35	59.5	56.1	36	3	AAW17306	AAW17306 TPO-mimet
36	59.5	56.1	36	3	AAW17306	AAW17306 TPO-mimet
37	59	55.7	18	5	ABP51688	ABP51688 TPO mimet
38	59	55.7	18	5	ABP51677	ABP51677 TPO mimet
39	59	55.7	18	5	ABP51675	ABP51675 TPO mimet
40	59	55.7	18	8	ADQ16641	ADQ16641 TPO mimet
41	59	55.7	18	8	ADQ16646	ADQ16646 TPO mimet
42	59	55.7	22	7	ADN59819	Adn59819 TWP pept1
43	59	55.7	128	8	ADQ16705	ADQ16705 Modified
44	59	55.7	225	8	ADQ16704	ADQ16704 Modified
45	59	55.7	472	5	ABP51695	ABP51695 5G1.1-TPO
46	59	55.7	472	8	ADQ16647	ADQ16647 Immunoglo
47	58	54.7	18	5	ABP51687	ABP51687 TPO mimet
48	58	54.7	18	5	ABP51693	ABP51693 TPO mimet
49	58	54.7	18	8	ADQ16617	ADQ16617 TPO mimet
50	58	54.7	18	8	ADQ16629	ADQ16629 TPO mimet
51	58	54.7	60	3	AAW17311	AAW17311 Synthetic
52	58	54.7	60	3	ABW73405	ABW73405 TWP-TWP G
53	58	54.7	247	5	AAW16961	AAW16961 TWP-Fc pr
54	58	54.7	247	5	ABW73414	ABW73414 TWP-Fc am
55	58	54.7	269	3	AAW16960	AAW16960 TWP-TWP-F
56	58	54.7	269	3	ABW73413	ABW73413 TWP-TWP-F
57	58	54.7	269	3	ABW73413	ABW73413 Immunoglo
58	57	53.8	18	5	ABP51686	ABP51686 TPO mimet
59	57	53.8	18	5	ABP51685	ABP51685 TPO mimet
60	57	53.8	18	7	ADN59663	Adn59663 Thrombopo
61	57	53.8	18	8	ADQ16615	ADQ16615 TPO mimet
62	57	53.8	18	8	ADQ16613	ADQ16613 TPO mimet
63	57	53.8	22	7	ADN59830	Adn59830 TWP pept1
64	57	53.8	25	7	ADN59708	Adn59708 Thrombopo
65	57	53.8	43	7	ADN59759	Adn59759 Peptide-v
66	56.5	53.3	36	3	AAW96523	AAW96523 Thrombopo
67	56.5	53.3	36	3	AAW17301	AAW17301 TPO-mimet
68	56.5	53.3	39	3	AAW17304	AAW17304 TPO-mimet
69	56.5	53.3	12	2	AAW36781	AAW36781 Thrombopo
70	56	52.8	13	2	AAW36779	AAW36779 Thrombopo
71	56	52.8	13	4	AAU26018	AAU26018 Human thr
72	56	52.8	13	4	AAU26008	AAU26008 Human thr
73	56	52.8	13	4	AAU26035	AAU26035 Human thr
74	56	52.8	13	8	ADW72525	Adm72525 TPO mimet
75	56	52.8	13	8	ADW72489	Adm72489 TPO mimet
76	56	52.8	13	8	ADW72488	Adm72488 TPO mimet
77	56	52.8	13	8	ADW72524	Adm72524 TPO mimet
78	56	52.8	14	2	AAW09468	AAW09468 Thrombopo
79	56	52.8	14	2	AAW09463	AAW09463 Thrombopo
80	56	52.8	14	2	AAW33030	AAW33030 Thrombopo
81	56	52.8	14	2	AAW36782	AAW36782 Thrombopo
82	56	52.8	14	2	AAW36774	AAW36774 Thrombopo
83	56	52.8	14	2	AAW66715	AAW66715 Peptide c
84	56	52.8	14	2	AAW66730	AAW66730 Peptide c
85	56	52.8	14	2	AD124843	AD124843 AF 12505
86	56	52.8	14	2	AAW96515	AAW96515 Thrombopo
87	56	52.8	14	3	AAW16962	AAW16962 TPO-mimet
88	56	52.8	14	3	AAW16962	AAW16962 TPO-mimet
89	56	52.8	14	3	AAW16968	AAW16968 TPO-mimet
90	56	52.8	14	4	AAU26009	AAU26009 Human thr
91	56	52.8	14	4	AAU26006	AAU26006 Human thr
92	56	52.8	14	4	AAU25827	AAU25827 Human thr
93	56	52.8	14	4	AAU26019	AAU26019 Human thr
94	56	52.8	14	4	AAU26036	AAU26036 Human thr
95	56	52.8	14	4	AAU26037	AAU26037 Human thr
96	56	52.8	14	4	AAU26004	AAU26004 Human thr
97	56	52.8	14	5	ABW72854	ABW72854 TPO mimet
98	56	52.8	14	5	ABW72853	ABW72853 TPO mimet

99 56 52.8 14 5 ABP51669  
100 56 52.8 14 5 AAE18011

Abp51669 Thrombopo  
AAE18011 Human lig

## ALIGNMENTS

RESULT 1  
AAW09494  
ID AAW09494 standard; protein; 19 AA.

AC AAW09494;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Haematology; thrombocytopenia; TPO; TR; proliferation;  
bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

PN WO9640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Matcheakls LC, Schatz PU, Magstrom CR, Wrighton NC;

DR WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 26; 106pp; English.

XX The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering

CC from haematological disorders and thrombocytopenia resulting from

CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide

CC may also be used to maintain the proliferation and growth of TPO-

CC dependent cell lines and for use in biological research, for detecting

CC TPO receptors on living cells

XX Sequence 19 AA;

XX Query Match 100.0%; Score 106; DB 2; Length 19;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-09;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHGNGRDT 19

DB 1 LAIEGPTLRQWLHGNGRDT 19

XX RESULT 2

XX AAW09461

XX ID AAW09461 standard; protein; 19 AA.

XX AC AAW09461;

XX XX

DT 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding compound peptide.

DE Thrombopoietin receptor binding compound peptide.

KW Haematology; thrombocytopenia; TPO; TR; proliferation;

KW bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 1.19

XX /note= "preferably linkages are selected from: -

CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6

FT -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is

FT lower alkyl"

FT Modified-site

FT /note= "preferably N-terminus is selected from: -NRR1; -

FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NRH; succinimide;

FT benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3

FT substitutions on the phenyl ring selected from lower

FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are

FT independently selected from hydrogen and lower alkyl"

FT Modified-site

FT /note= "preferably C-terminus is -C(O)R2 where R2 is

FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3

FT and R4 are independently selected from hydrogen and lower

FT alkyl, and where the nitrogen atom of the -NR3R4 group

FT can optionally be the amine group of the N-terminus of

FT the peptide forming a cyclic peptide"

XX WO9640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

XX 07-JUN-1995; 95US-00473604.

XX 07-JUN-1995; 95US-00476168.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00484090.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Matcheakls LC, Schatz PU, Magstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

PS Claim 18; Page 89; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)

CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding

CC affinity to TR as expressed by an IC50 of no more than about 100 nM. The

CC compound (especially if modified, see features table) can be used for

CC treating patients suffering from haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. The peptide may also be used to maintain the

CC proliferation and growth of TPO-dependent cell lines and for use in

CC biological research, for detecting TPO receptors on living cells

XX Sequence 19 AA;

XX Query Match 100.0%; Score 106; DB 2; Length 19;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-09;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHGNGRDT 19



Db 1 LAIEGPTLRQWLHNGRDT 19

RESULT 3  
AAW36645 standard; peptide; 19 AA.

AAW36645;

11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;  
haematological disorder; thrombocytopenia; chemotherapy;  
radiation therapy; bone marrow transfusion; diagnosis;  
signal transduction; receptor activation; cell culture.

Synthetic.

WO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAXO) GLAXO GROUP LTD.

Dower WJ, Barrett RM, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;  
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the  
thrombopoietin receptor - useful in treatment of haematological  
disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
Disclosure; Page 26; 106pp; English.

The present peptide, which binds the thrombopoietin receptor (TR), can be  
used to treat disorders which are susceptible to treatment with a  
thrombopoietin agonist, preferably haematological disorders and  
thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
marrow transfusions. It can also be used diagnostically, e.g. to  
investigate the mechanism of thrombopoietin signal transduction and  
receptor activation, or to maintain the proliferation and growth of  
thrombopoietin dependent cell lines

Sequence 19 AA;

Query Match 100.0%; Score 106; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHNGRDT 19  
DB 1 LAIEGPTLRQWLHNGRDT 19

RESULT 4

AAAB17022 standard; peptide; 19 AA.

AAAB17022;

31-OCT-2000 (first entry)

TPO-mimetic peptide sequence SEQ ID NO:78.

Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
vascular endothelial growth factor; matrix metalloproteinase; asthma;  
thrombosis; pharmaceutical.

Synthetic.

WO200024782-A2.

04-MAY-2000.

25-OCT-1999; 99WO-US025044.

23-OCT-1998; 98US-0105371P.

22-OCT-1999; 99US-00428082.

(AMGE-) AMGEN INC.

Feige U, Liu C, Cheetham J, Boone TC;

WPI; 2000-350702/30.

Novel composition of matter comprising an Fc domain and pharmacologically  
active peptides, useful for treating cancer and autoimmune diseases.  
Claim 19; Page 221; 608pp; English.

The present invention describes composition of matter (I) comprising an  
Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
(X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-  
(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,  
P3, and P4 = are each independently sequences of pharmacologically active  
peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
c, d, e, and f = are each independently 0 or 1, provided that at least 1  
of a and b is 1. The composition can have cytostatic, antiasthmatic,  
thrombolytic and immunosuppressive activities. DNAs, vectors and host  
cells from the present invention can be used for producing pharmaceutical  
compositions. The compositions are useful for treating cancer, asthma,  
thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
a Fab domain) can provide a longer half-life or incorporate functions  
such as Fc receptor binding, protein A binding, complement fixation, and  
possibly placental transfer. AA69443 to AA69526 and AA61955 to  
AA61903 represent nucleotide and amino acid sequences used in the  
exemplification of the present invention

Sequence 19 AA;

Query Match 100.0%; Score 106; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHNGRDT 19  
DB 1 LAIEGPTLRQWLHNGRDT 19

RESULT 5

AAU25864 standard; peptide; 19 AA.

AAU25864;

17-DEC-2001 (first entry)

Human thrombopoietin receptor (TPO-R) activator peptide #50.

Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
haemostatic; thrombocytopenia; chemotherapy; radiation therapy; EritrA;  
bone marrow transplantation; haematological disorder; platelet disorder;

KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS Homo sapiens.

PN US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009622.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balaubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddurturi S;  
 PI Yin Q;

DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 20; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 19 AA:

Query Match 100.0%; Score 106; DB 4; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.4e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHGNGRDT 19

DB 1 LAIEGPTLRQWLHGNGRDT 19

RESULT 6

AAU25825

XX AAU25825 standard; peptide; 19 AA.

AC AAU25825;

XX 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #11.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

KW bone marrow transplantation; hematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS Homo sapiens.

PN US6251864-B1.

XX 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009622.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balaubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddurturi S;  
 PI Yin Q;

DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 67-68; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 19 AA:

Query Match 100.0%; Score 106; DB 4; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.4e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHGNGRDT 19

DB 1 LAIEGPTLRQWLHGNGRDT 19

RESULT 7

ABR72908

XX ABR72908 standard; peptide; 19 AA.

AC ABR72908;

XX 05-APR-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:78.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;

KM erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KM TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KM TPO mimetic peptide; EPO mimetic peptide; EGF; VEGF antagonist;  
 KM MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KM cytostatic; antineoplastic; antiferility; haemostatic; dermatological;  
 KM antianemic; anorectic; antiferility; haemostatic; dermatological;  
 KM neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KM cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KM sleep disorder; neurological degenerative disease; anaemia;  
 KM thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 KM Fanconi's syndrome.  
 XX Homo sapiens.  
 OS Synthetic.  
 PN WO200183525-A2.  
 PD 08-NOV-2001.  
 PF 02-MAY-2001; 2001WO-US014310.  
 PR 03-MAY-2000; 2000US-00563286.  
 XX (AMGEN-) AMGEN INC.  
 PI Feige U, Liu C, Cheetham JC, Boone TC, Gudae JM;  
 DR WPI; 2002-130313/17.  
 XX Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 44; 176pp; English.  
 XX The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antineoplastic, antiferility, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antiferility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB5565 to ABB5777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 19 AA:  
 Query Match 100.0%; Score 106; DB 5; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LAIEGPTLRQWLHNGRDT 19  
 DB 1 LAIEGPTLRQWLHNGRDT 19  
 RESULT 8  
 ADJ73060 standard; peptide; 19 AA.  
 ID ADJ73060 standard; peptide; 19 AA.  
 XX

AC ADJ73060;  
 XX 06-MAY-2004 (first entry)  
 DT TPO mimetic peptide sequence SeqID 514.  
 XX  
 DE TPO mimetic peptide sequence SeqID 514.  
 XX  
 KM mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KM immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
 KM TPO.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2003084477-A2.  
 PD 16-OCT-2003.  
 PF 24-MAR-2003; 2003WO-US009139.  
 PR 29-MAR-2002; 2002US-0368791P.  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;  
 DR WPI; 2003-804237/75.  
 XX  
 PT New CDR mimetibody comprising a portion of a heavy or light chain  
 PT variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PT cardiovascular or neurologic disease.  
 XX  
 PS Disclosure; SEQ ID NO 514; 97pp; English.  
 XX  
 CC This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LSR). The present invention describes human  
 CC mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neurologic  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 XX  
 SQ Sequence 19 AA:  
 Query Match 100.0%; Score 106; DB 7; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LAIEGPTLRQWLHNGRDT 19  
 DB 1 LAIEGPTLRQWLHNGRDT 19  
 RESULT 9  
 ADJ52695 standard; peptide; 19 AA.  
 ID ADJ52695 standard; peptide; 19 AA.  
 AC ADJ52695;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CH1 deleted mimetibody-related peptide SeqID514.  
 XX  
 KM CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arrhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 XX WO2004002417-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 XX 27-JUN-2003; 2003WO-US020347.  
 PF  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX Heavenr GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutoloeki KA;  
 XX  
 DR WPI; 2004-082870/08.  
 XX  
 PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 XX  
 PS Claim 2; SEQ ID NO 514; 129pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 CC  
 SQ Sequence 19 AA;  
 XX  
 QY Query Match 100.0%; Score 106; DB 8; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 1 LAIEGPTLRQWLHNGRDT 19  
 1 LAIEGPTLRQWLHNGRDT 19  
 XX  
 RESULT 10  
 ADJ51656  
 ID ADJ51656 standard; peptide; 19 AA.  
 XX  
 AC ADJ51656;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CHI deleted mimetibody-related peptide SegID514.  
 XX  
 KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;  
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;

KW ophthalmological; nephrotoxic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstetric disorder; haematologic disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 XX WO2004002424-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 XX 30-JUN-2003; 2003WO-US020495.  
 PF  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX Heavenr GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutoloeki KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 15; SEQ ID NO 514; 123pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, anti-allergic, muscular-Gen, cytostatic,  
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotoxic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 SQ Sequence 19 AA;  
 XX  
 QY Query Match 100.0%; Score 106; DB 8; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 1 LAIEGPTLRQWLHNGRDT 19  
 1 LAIEGPTLRQWLHNGRDT 19  
 XX  
 RESULT 11  
 AAW33028  
 ID AAW33028 standard; peptide; 19 AA.  
 XX  
 AC AAW33028;





XX	AAB17285
ID	AAB17285 standard; peptide; 28 AA.
AC	
XX	AAB17285;
DT	31-OCT-2000 (first entry)
XX	
DE	TPO-mimetic peptide sequence SEQ ID NO:341.
XX	
KM	Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW	autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW	immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KM	inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW	cytotoxic T cell lymphocyte antigen 4; tumor necrosis factor;
KM	vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW	thrombolysis; pharmaceutical.
XX	
OS	Synthetic.
XX	
PN	WO200024782-A2.
XX	
PD	04-MAY-2000.
XX	
PF	25-OCT-1999; 99WO-US025044.
XX	
PR	23-OCT-1998; 98US-0105371P.
XX	
PR	22-OCT-1999; 99US-00428082.
XX	
PA	(AMGE-) AMGEN INC.
XX	
PI	Feige U, Liu C, Cheetham J, Boone TC;
XX	
WP	WI; 2000-350702/30.
XX	
PT	Novel composition of matter comprising an Fc domain and pharmacologically
PR	active peptides, useful for treating cancer and autoimmune diseases.
XX	
XX	Example 1; Page 315; 608pp; English.
PS	
XX	
CC	The present invention describes composition of matter (I) comprising an
CC	Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC	(X1)-a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC	independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC	(L2)d-P2-(L3)e-F3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC	P3, and P4 = are each independently sequences of pharmacologically active
CC	peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC	c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC	of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC	thrombolytic, and immunosuppressive activities. DNAs, vectors and host
CC	cells from the present invention can be used for producing pharmaceutical
CC	compositions. The compositions are useful for treating cancer, asthma,
CC	chromobias, or autoimmune diseases. The use of an Fc domain (rather than
CC	a Fab domain) can provide a longer half-life or incorporate functions
CC	such as Fc receptor binding, protein A binding, complement fixation, and
CC	possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC	AAB18003 represent nucleotide and amino acid sequences used in the
CC	embodiment of the present invention
XX	
SQ	Sequence 28 AA;
XX	
Query Match	56.6%; Score 60; DB 3; Length 28;
Best Local Similarity	100.0%; Pred. No. 0.04;
Matches 11; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	2 AIEGPTLRQWL 12 
Db	14 AIEGPTLRQWL 24
RESULT 16	
ADJ73011	
ID	ADJ73011 standard; peptide; 29 AA.
XX	

AC	ADJ73011;
DT	06-MAY-2004 (first entry)
DE	TPO mimetic peptide sequence SegID 465.
KM	mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW	cardiovascular; infectious; malignant; neurologic disease; anaemia;
KX	immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
XX	TPO.
XX	Synthetic.
XX	
PN	WO2003084477-A2.
PD	16-OCT-2003.
PF	24-MAR-2003; 2003WO-US009139.
PR	29-MAR-2002; 2002US-0368791P.
PA	(CENZ ) CENTOCOR INC.
PI	Heavner GA, Knight DW, Scallion BJ, Ghrayeb J;
PT	WPI; 2003-804237/75.
PT	New CDR mimetibody comprising a portion of a heavy or light chain
PT	variable region comprising human framework or ligand binding region,
PT	useful for preparing a composition for treating e.g., immune,
PT	cardiovascular or neurologic disease.
PS	
PS	Disclosure; SEQ ID NO 465; 97bp; English.
CC	This invention relates to novel mammalian CDR mimetibodies, specific
CC	portions or variants thereof. Specifically, it refers to an antibody
CC	fragment where a protein has been inserted into, or replaces a portion
CC	of, one or more CDR regions, such that each CDR mimetibody comprises at
CC	least one portion of a heavy chain or light chain variable region, which
CC	itself comprises at least one human framework region and at least one
CC	ligand binding region (LBR). The present invention describes human
CC	mimetibodies, including modified immunoglobulins and cleavage products
CC	that can be useful in gene therapy and the generation of transgenic
CC	plants and animals. Furthermore, the CDR mimetibody is useful for
CC	preparing compositions for modulating, treating or reducing the symptoms
CC	of immune, cardiovascular, infectious, malignant and/or neurologic
CC	diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC	cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC	peptide sequence is a TPO mimetic peptide sequence used to make a
CC	mimetibody of the invention.
XX	
XX	
SQ	Sequence 29 AA;
Query Match	56.6%; Score 60; DB 7; Length 29;
Best Local Similarity	100.0%; Pred. No. 0.042;
Matches 11; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	2 AIEGPTLRQWL 12       
DB	15 AIEGPTLRQWL 25
RESULT 17	
ADJ73007	
ID	ADJ73007 standard; peptide; 29 AA.
XX	
XX	ADJ73007;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	TPO mimetic peptide sequence SegID 461.
XX	
KM	mimetic; CDR mimetibody; gene therapy; transgenic; immune;







PF 27-JUN-2003; 2003WO-US020347.  
 XX  
 XX 28-JUN-2002; 2002US-0392431P.  
 XX  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX  
 PI Heavenr GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutooski KA;  
 XX  
 DR WPI; 2004-082870/08.  
 XX  
 PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 PS  
 XX Claim 2; SEQ ID NO 461; 129pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 SX  
 SQ Sequence 29 AA;  
 QY  
 Query Match 56.6%; Score 60; DB 8; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.042;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 2 AIEGPTLRQWL 12  
 15 AIEGPTLRQWL 25  
 QY  
 RESULT 20  
 ADJ52646  
 ID ADJ52646 standard; peptide; 29 AA.  
 AC  
 XX ADJ52646;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CHI deleted mimetibody-related peptide SeqID465.  
 XX  
 XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arrhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 XX  
 PN WO2004002417-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 27-JUN-2003; 2003WO-US020347.  
 XX  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR

XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 XX Heavenr GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutooski KA;  
 XX  
 DR WPI; 2004-082870/08.  
 XX  
 PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 PS  
 XX Claim 2; SEQ ID NO 465; 129pp; English.  
 XX  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 SX  
 SQ Sequence 29 AA;  
 QY  
 Query Match 56.6%; Score 60; DB 8; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.042;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 2 AIEGPTLRQWL 12  
 15 AIEGPTLRQWL 25  
 QY  
 RESULT 21  
 ADJ52641  
 ID ADJ52641 standard; peptide; 29 AA.  
 AC  
 XX ADJ52641;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CHI deleted mimetibody-related peptide SeqID460.  
 XX  
 XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arrhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 XX  
 PN WO2004002417-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 27-JUN-2003; 2003WO-US020347.  
 XX  
 XX 28-JUN-2002; 2002US-0392431P.  
 PA (CENZ ) CENTOCOR INC.  
 XX

PR 28-JUN-2002; 2002US-0392433.P.  
PR 19-SEP-2002; 2002US-0412144P.  
XX  
XX  
PA (CENZ ) CENTOCOR INC.  
XX  
FI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
P1 Kutolooski KA;  
XX  
XX  
DR WPI; 2004-082872/08.  
XX  
PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
XX

XX XX  
PR 28-JUN-2002; 2002US-0392431P.  
FR 19-SEP-2002; 2002US-0412144P.

PA (CENZ ) CENTOCOR INC.  
XX  
FI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;  
PI Kutolooski KA;

DR WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
PT diagnosing, preventing or treating cardiovascular, dermatologic, and  
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
PT nutritional disorders.

XX  
PS Claim 14; SEQ ID NO 461; 123pp; English.

CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an osteopathic,  
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
CC gastrointestinal-Gen, gynaeecological-Gen, hepatotropic, haemostatic,  
CC immunomodulator, antiallergic, muscular-Gen, cytotoxic,  
CC antiinflammatory, neuroleptic, ophthalmological, nephrotoxic or  
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
CC modulator or cytokine-agonist. The methods and compositions of the  
CC present invention are useful for the diagnosis, prevention and/or  
CC treatment of diseases or conditions associated with aberrant expression  
or activity of the CHI deleted mimetibody, such as a bone or joint,

CC endocrine, hematologic, gastrointestinal, gynecological, hepatic,  
CC obstructive, rheumatologic, immunological, allergic, infectious,  
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
CC pediatric, psychiatric, renal or pulmonary disorder. The present  
CC sequence is that of a peptide which may be used during the creation of a  
CC mimetbody of the invention.  
SQ Sequence 29 AA;  
XX

Query Match	56.6%;	Score	60;	DB	8;	Length	29;
Best Local Similarity	100.0%;	Pred. No.	0.042;				
Matches	11;	Conservative	0;	Mismatches	0;	Indels	0;
Qy	2	AIEGPTLRWL	12				
Db	15	AIEGPTLRWL	25				

ADJ51602	
ID	ADJ51602 standard; peptide; 29 AA.
XX	
AC	ADJ51602;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	CHI deleted mimeribody-related peptide SegID#60.
XX	
KW	CHI deleted mimeribody; osteopathic; cardiovascular-Gen;
KW	dematological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW	gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW	antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW	ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor

KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
KW obstetric disorder; haematological disorder; immunological disorder;  
KW allergic disorder; infectious disorder; musculoskeletal disorder;  
KW oncological disorder; neurological disorder; nutritional disorder;  
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
KW renal disorder; pulmonary disorder.

OS Unidentified.  
 OS Synthetic.  
 XX WO2004002424-A2.  
 XX  
 XX 08-JAN-2004.  
 PD  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutolowski KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 14; SEQ ID NO 460; 123pp; English.  
 XX  
 XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CH1 deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstructive, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncologic, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 SQ Sequence 29 AA;  
 XX  
 XX Query Match 56.6%; Score 60; DB 8; Length 29;  
 XX Best Local Similarity 100.0%; Pred. No. 0.042;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AIEGPTLRQWL 12  
 XX |||||||||  
 DB 15 AIEGPTLRQWL 25  
 XX  
 XX RESULT 24  
 XX ADJ51607 standard; peptide; 29 AA.  
 ID ADJ51607 standard; peptide; 29 AA.  
 XX  
 XX ADJ51607;  
 AC  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE CH1 deleted mimetibody-related peptide SeqID465.  
 XX  
 XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;  
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstructive disorder; haematologic disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 XX  
 XX Unidentified.  
 OS Synthetic.  
 XX WO2004002424-A2.  
 XX  
 XX 08-JAN-2004.  
 PD  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 XX 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;  
 PI Kutolowski KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 14; SEQ ID NO 465; 123pp; English.  
 XX  
 XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,  
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CH1 deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstructive, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncologic, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 SQ Sequence 29 AA;  
 XX  
 XX Query Match 56.6%; Score 60; DB 8; Length 29;  
 XX Best Local Similarity 100.0%; Pred. No. 0.042;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AIEGPTLRQWL 12  
 XX |||||||||  
 DB 15 AIEGPTLRQWL 25  
 XX  
 XX RESULT 25  
 XX ADJ73009 standard; peptide; 31 AA.  
 ID ADJ73009 standard; peptide; 31 AA.  
 XX  
 XX ADJ73009;  
 AC  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX



Key Location/Qualifiers  
 Misc-difference 16 /note="Residue is a BrAc residue"

WO2004002417-A2.  
 08-JAN-2004.  
 27-JUN-2003; 2003WO-US020347.  
 28-JUN-2002; 2002US-0392431P.  
 (CENZ ) CENTOCOR INC.  
 Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC,  
 Kutoloski KA;  
 WPI; 2004-082870/08.

New CH1-deleted mimetibody polypeptides and nucleic acids, useful for modulating, treating, alleviating, preventing an immune, cardiovascular, or neurodegenerative disease or disorder, anemia, cancer, or infectious diseases.

Claim 2; SEQ ID NO 463; 129pp; English.

This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an immunosuppressive, cardiovascular, cardiant, hypotensive, neuroprotective, nootropic, antibacterial, virucide or fungicide activity. In addition, the disclosed sequences may prove useful for gene therapy. The CH1-deleted mimetibody is useful for diagnosing or treating a disease condition in a cell, tissue, organ or animal, specifically for modulating, treating, alleviating, preventing the incidence or reducing the symptoms of an immune, cardiovascular (for example arrhythmia, hypertension or heart failure), or neurodegenerative (for example multiple sclerosis, dementia or Alzheimer's disease) diseases or disorders, anaemia, cancerous conditions, or infectious diseases (for example bacterial, viral or fungal infection). The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.

Sequence 31 AA;

Query Match 56.6%; Score 60; DB 8; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 AIEGPTLRQWL 12  
 17 AIEGPTLRQWL 27

RESULT 28  
 ADJ52645  
 ID ADJ52645 standard; peptide; 31 AA.  
 AC ADJ52645;  
 DT 06-MAY-2004 (first entry)  
 DE CH1 deleted mimetibody-related peptide SeqID464.  
 XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;  
 XX hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 XX fungicide; gene therapy; immune disorder; cardiovascular disease;  
 XX arrhythmia; hypertension; heart failure; neurodegenerative;  
 XX multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 XX cancerous condition; infectious diseases; bacterial infection;  
 XX viral infection; fungal infection.  
 XX Undifferentiated.  
 OS Synthetic.

Key Location/Qualifiers  
 Misc-difference 16 /note="Residue is a PEG residue"

WO2004002417-A2.  
 08-JAN-2004.  
 27-JUN-2003; 2003WO-US020347.  
 28-JUN-2002; 2002US-0392431P.  
 (CENZ ) CENTOCOR INC.  
 Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC,  
 Kutoloski KA;  
 WPI; 2004-082870/08.

New CH1-deleted mimetibody polypeptides and nucleic acids, useful for modulating, treating, alleviating, preventing an immune, cardiovascular, or neurodegenerative disease or disorder, anemia, cancer, or infectious diseases.

Claim 2; SEQ ID NO 464; 129pp; English.

This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an immunosuppressive, cardiovascular, cardiant, hypotensive, neuroprotective, nootropic, antibacterial, virucide or fungicide activity. In addition, the disclosed sequences may prove useful for gene therapy. The CH1-deleted mimetibody is useful for diagnosing or treating a disease condition in a cell, tissue, organ or animal, specifically for modulating, treating, alleviating, preventing the incidence or reducing the symptoms of an immune, cardiovascular (for example arrhythmia, hypertension or heart failure), or neurodegenerative (for example multiple sclerosis, dementia or Alzheimer's disease) diseases or disorders, anaemia, cancerous conditions, or infectious diseases (for example bacterial, viral or fungal infection). The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.

Sequence 31 AA;

Query Match 56.6%; Score 60; DB 8; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 AIEGPTLRQWL 12  
 17 AIEGPTLRQWL 27

RESULT 29  
 ADJ51606  
 ID ADJ51606 standard; peptide; 31 AA.  
 AC ADJ51606;  
 DT 06-MAY-2004 (first entry)  
 DE CH1 deleted mimetibody-related peptide SeqID464.  
 XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
 XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 XX gynaecological-Gen; hepatotropic; haemostatic; immunomodulatory;  
 XX antiallergic; muscular-Gen; cytostatic; antiinflammatory; neurologic;  
 XX ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
 XX TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 XX dental disorder; oral disorder; dermatological disorder; ear disorder;  
 XX nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 XX gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KM obstetric disorder; haematologic disorder; immunological disorder;  
 KM allergic disorder; infectious disorder; musculoskeletal disorder;  
 KM oncological disorder; neurological disorder; nutritional disorder;  
 KM ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KM renal disorder; pulmonary disorder.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 FT Key Location/Qualifiers  
 FT Misc-difference 16  
 FT note="Residue is a PEG residue"  
 XX  
 PN WO2004002424-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesper TC;  
 PI Kutoloeki KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 14; SEQ ID NO 464; 123pp; English.  
 CC  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, anti-allergic, muscular-Gen, cyostatic,  
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 CC Sequence 31 AA;  
 SQ  
 Query Match 56.6%; Score 60; DB 8; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AIBGPTLRQWL 12  
 DB 17 AIBGPTLRQWL 27  
 RESULT 30  
 ADJ51605  
 ID ADJ51605 standard; peptide; 31 AA.  
 XX  
 AC ADJ51605;  
 XX

DT 06-MAY-2004 (first entry)  
 XX  
 DE CHI deleted mimetibody-related peptide Segid463.  
 XX  
 KM CHI deleted mimetibody; osteopathic; cardiovascular-Gen;  
 KM dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
 KM gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KM anti-allergic; muscular-Gen; cyostatic; anti-inflammatory; neuroleptic;  
 KM ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
 KM TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KM dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KM nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KM gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KM obstetric disorder; haematologic disorder; immunological disorder;  
 KM allergic disorder; infectious disorder; musculoskeletal disorder;  
 KM oncological disorder; neurological disorder; nutritional disorder;  
 KM ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
 KM renal disorder; pulmonary disorder.  
 XX  
 OS Unidentified.  
 OS Synthetic.  
 FT Key Location/Qualifiers  
 FT Misc-difference 16  
 FT note="Residue is a BrAc residue"  
 XX  
 PN WO2004002424-A2.  
 XX  
 PD 08-JAN-2004.  
 XX  
 PF 30-JUN-2003; 2003WO-US020495.  
 XX  
 PR 28-JUN-2002; 2002US-0392431P.  
 PR 19-SEP-2002; 2002US-0412144P.  
 XX  
 PA (CENZ ) CENTOCOR INC.  
 XX  
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesper TC;  
 PI Kutoloeki KA;  
 XX  
 DR WPI; 2004-082872/08.  
 XX  
 PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic,  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX  
 PS Claim 14; SEQ ID NO 463; 123pp; English.  
 CC  
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, anti-allergic, muscular-Gen, cyostatic,  
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or  
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 CC  
 CC Sequence 31 AA;  
 SQ  
 Query Match 56.6%; Score 60; DB 8; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AIBGPTLRQWL 12  
 DB 17 AIBGPTLRQWL 27  
 RESULT 30  
 ADJ51605  
 ID ADJ51605 standard; peptide; 31 AA.  
 XX  
 AC ADJ51605;  
 XX

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AIEGPTLRQWL 12  
17 AIEGPTLRQWL 27

RESULT 31  
ADN59659  
ID ADN59659 standard; peptide; 18 AA.

AC ADN59659;

DT 01-JUL-2004 (first entry)

DE Thrombopoietin mimetic peptide (TMP8), seq id 8.

XX Haemostatic; antihaemic; immunosuppressive; platelet;  
XX transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
XX TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
XX thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
XX autoimmune haemolytic anaemia; Hughes's syndrome;  
XX lupoid thrombocytopenia.

OS Homo sapiens.

PN WO2003031589-A2.

PD 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

XX Min H, Sitney KC, Hartley C;

PI MPI; 2003-403101/38.

PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
PT which stimulate the production of platelets and/or the production of  
PT platelet precursors, useful for treating thrombocytopenia.

PS Claim 6; SEQ ID NO 8; 126pp; English.

CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
CC platelets and/or the production of platelet precursors, is new. Further  
CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
CC and a pharmaceutical composition comprising (II) and a carrier. The  
CC pharmaceutical composition of the invention is useful for treating  
CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
CC platelets in a patient. The TMP of the invention is useful for treating  
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
CC thrombocytopenia. The TMP of the invention is also useful for  
CC maintaining the viability or storage life of platelets and/or  
CC megakaryocytes and its derived cells. The compounds demonstrate an  
CC improved ability to bind to and/or trigger transmembrane signal through,  
CC i.e. activating, the mpl receptor the compounds have superior  
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
CC vitro, the production of platelets and/or megakaryocytopenic activity,  
CC i.e. the ability to stimulate, in vivo and in vitro, the production of,  
CC platelet precursors. Further, certain of the compounds also exhibit  
CC biological activity and in vivo circulation time. The current sequence  
CC represents a preferred TMP of the invention.

XX Sequence 18 AA;

Query Match 56.1%; Score 59.5; DB 7; Length 18;  
Best Local Similarity 73.3%; Pred. No. 0.029;  
Matches 11; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 4 EGPTLRQWL-HNGR 17  
4 EGPTLRQWL-HNGRQ 18

RESULT 32  
ADN59826  
ID ADN59826 standard; peptide; 22 AA.

AC ADN59826;

DT 01-JUL-2004 (first entry)

DE TMP peptide TMP8.

XX Haemostatic; antihaemic; immunosuppressive; platelet;  
XX transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
XX TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
XX thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;  
XX autoimmune haemolytic anaemia; Hughes's syndrome;  
XX lupoid thrombocytopenia; linker.

OS Homo sapiens.

PN WO2003031589-A2.

PD 17-APR-2003.

PF 11-OCT-2002; 2002WO-US032552.

PR 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

XX Min H, Sitney KC, Hartley C;

PI MPI; 2003-403101/38.

PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
PT which stimulate the production of platelets and/or the production of  
PT platelet precursors, useful for treating thrombocytopenia.

PS Example 6; Page 83; 126pp; English.

CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that  
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of  
CC platelets and/or the production of platelet precursors, is new. Further  
CC disclosed is a composition of matter (II) that binds to an mpl receptor,  
CC and a pharmaceutical composition comprising (II) and a carrier. The  
CC pharmaceutical composition of the invention is useful for treating  
CC thrombocytopenia in an animal, and for increasing megakaryocytes or  
CC platelets in a patient. The TMP of the invention is useful for treating  
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.,  
CC disease conditions involving thrombocytopenia such as aplastic anaemia,  
CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,  
CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid  
CC thrombocytopenia. The TMP of the invention is also useful for  
CC maintaining the viability or storage life of platelets and/or  
CC megakaryocytes and its derived cells. The compounds demonstrate an  
CC improved ability to bind to and/or trigger transmembrane signal through,  
CC i.e. activating, the mpl receptor the compounds have superior  
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in  
CC vitro, the production of platelets and/or megakaryocytopenic activity,  
CC i.e. the ability to stimulate, in vivo and in vitro, the production of,  
CC platelet precursors. Further, certain of the compounds also exhibit  
CC biological activity and in vivo circulation time. The current sequence  
CC represents a preferred TMP of the invention.



CC represents a TMP peptide of the invention to which a two amino acid "cap"  
 CC has been added to the carboxy terminal to increase peptide affinity.  
 XX  
 XX  
 SQ Sequence 22 AA;

Query Match 56.1%; Score 59.5; DB 7; Length 22;  
 Best Local Similarity 73.3%; Pred. No. 0.037;  
 Matches 11; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 4 EGPTRLRQWL-HGNGR 17  
 |||||:|||||:  
 Db 6 EGPTRLKQWLHFGRGQ 20

## RESULT 33

ADNS9700  
 ID ADNS9700 standard; peptide; 25 AA.

AC ADNS9700;

XX 01-JUL-2004 (first entry)

DT Thrombopoietin mimetic peptide TMP8, seq id 49.

DE Haemostatic; antihaemic; immunosuppressive; platelet;

XX transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;

KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;

KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;

KW autoimmune haemolytic anaemia; Hughes's syndrome;

XX lupoid thrombocytopenia.

OS Homo sapiens.

XX WO2003031589-A2.

PN 17-APR-2003.

XX 11-OCT-2002; 2002WO-US032552.

XX 11-OCT-2001; 2001US-0328666P.

PR 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

PA Min H, Sicney KC, Hartley C;

PI WPI; 2003-403101/38.

XX N-PSDB; ADNS9699.

DR Novel thrombopoietin mimetic peptides which bind to mpl receptor, and

XX which stimulate the production of platelets and/or the production of

PT platelet precursors, useful for treating thrombocytopenia.

PS Disclosure; SEQ ID NO 49; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that

CC binds to the c-mpl (mpl) receptor, and which stimulates the production of

CC platelets and/or the production of platelet precursors, is new. Further

XX disclosed is a composition of matter (II) that binds to an mpl receptor,

CC and a pharmaceutical composition comprising (II) and a carrier. The

XX pharmaceutical composition of the invention is useful for treating

CC thrombocytopenia in an animal, and for increasing megakaryocytes or

XX platelets in a patient. The TMP of the invention is useful for treating

CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.

CC disease conditions involving thrombocytopenia such as aplastic anaemia,

CC autoimmune thrombocytopenia, drug induced immune thrombocytopenia,

CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid

CC thrombocytopenia. The TMP of the invention is also useful for

CC maintaining the viability or storage life of platelets and/or

CC megakaryocytes and its derived cells. The compounds demonstrate an

CC improved ability to bind to and/or trigger transmembrane signal through,

CC i.e. activating, the mpl receptor the compounds have superior

CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in

CC vitro, the production of platelets and/or megakaryocytopenic activity,  
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of  
 CC platelet precursors. Further, certain of the compounds also exhibit  
 CC superior therapeutic properties, such as improved plasma half-life.  
 CC CC biological activity and in vivo circulation time. The current sequence  
 CC represents a TMP fragment.

XX Sequence 25 AA;

Query Match 56.1%; Score 59.5; DB 7; Length 25;  
 Best Local Similarity 73.3%; Pred. No. 0.042;  
 Matches 11; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 4 EGPTRLRQWL-HGNGR 17  
 |||||:|||||:  
 Db 7 EGPTRLKQWLHFGRGQ 21

## RESULT 34

AA96526  
 ID AA96526 standard; peptide; 36 AA.

XX AA96526;

XX 04-SEP-2000 (first entry)

DT Thrombopoietin mimetic peptide compound 7.

DE Thrombopoietin; mimetic; TMP; TPO; platelet; megakaryocyte; production;

KW anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;

KW immunosuppressive; anti-inflammatory; linker.

XX Synthetic.

XX Key

XX Location/Qualifiers

XX Peptide

XX Modified-site

XX 1

XX /note= "optionally linked to an Fc molecule"

XX Peptide

XX /label= linker

XX Peptide

XX /label= TMP\_2

XX MO200024770-A2.

XX 04-MAY-2000.

XX 22-OCT-1999; 99WO-US024834.

XX 23-OCT-1998; 98US-0105348P.

XX (AMGE-) AMGEN INC.

XX Liu C, Feige U, Cheatham J;

XX WPI; 2000-365108/31.

XX Thrombopoietic peptides which activate mpl receptors and increase the

XX production of platelets or platelet precursors, useful for treatment of

XX diseases which involve thrombocytopenia.

XX Claim 16; Page 62; 91pp; English.

XX A compound which binds to an mpl receptor comprising a thrombopoietin

XX mimetic peptide (TMP) dimer joined by a linker (TMP\_1-(L\_1)-TMP\_2), is

XX new. TMP 1 and TMP 2 are amino acid sequences varying from at least 10 to

XX 14 residues in length comprising X<sub>2</sub>-X<sub>1</sub> 0, X<sub>2</sub>-X<sub>1</sub> 1, X<sub>2</sub>-X<sub>1</sub> 2, X<sub>2</sub>-

XX X<sub>1</sub> 3, X<sub>2</sub>-X<sub>1</sub> 4, X<sub>1</sub>-X<sub>1</sub> 0, X<sub>1</sub>-X<sub>1</sub> 1, X<sub>1</sub>-X<sub>1</sub> 2, X<sub>1</sub>-X<sub>1</sub> 3 and X<sub>1</sub>-

XX X<sub>1</sub> 4. X<sub>1</sub>=I, A, V, L, S or R; X<sub>2</sub>=E, D, K or V; X<sub>3</sub>=G or A; X<sub>4</sub>=

XX F; X<sub>5</sub>=T or S; X<sub>6</sub>=L, I, V, A, F, M, or K; X<sub>7</sub>=R or K; X<sub>8</sub>=Q, N, or E;

XX X<sub>9</sub>=W, Y or F; X<sub>10</sub>=L, I, V, A, F, M, or K; X<sub>11</sub>=A, I, V, L, F,

XX S, T, K, H, or E; X<sub>12</sub>=A, I, V, L, F, G, S, or Q; X<sub>13</sub>=R, K, T, V,



CC N, Q or G; X<sub>1-4</sub> = A, I, V, L, F, T, R, E, or G; L<sub>1</sub> = linker comprising  
 CC 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate  
 CC the c-Mpl receptor which mediates the activity of endogenous  
 CC thrombopoietin. The TMs are useful for increasing the production of  
 CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which  
 CC is useful for treatment of diseases which involve thrombocytopenia, e.g.  
 CC aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency  
 CC virus associated ITP, and systemic lupus erythematosus

XX Sequence 36 AA;

Query Match 56.1%; Score 59.5; DB 3; Length 36;

Best Local Similarity 68.4%; Pred. No. 0.064; 1; Indels 5; Gaps 1;

Matches 13; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

3 IEGPTLRQWL-----HGNG 16  
 1 IEGPTLRQWLAAAGGNG 19

Db

RESULT 35

AAB17306 standard; peptide; 36 AA.

XX AAB17306;

XX 31-OCT-2000 (first entry)

XX TPO-mimetic peptide sequence SEQ ID NO:362.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 XX autoimmune disease; cytostatic; antineoplastic; thrombolytic; VEGF;  
 XX immunosuppressive; EPO; TPO; CTAP4; mimetic; IL-1; TNF; antagonist; MMP;  
 XX inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 XX cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 XX vascular endothelial growth factor; matrix metalloproteinase; aschma;  
 XX thrombosis; pharmaceutical.

XX Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Peige U, Liu C, Cheetham J, Boone TC,

XX WPI, 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically  
 XX active peptides, useful for treating cancer and autoimmune diseases.

XX Example 1; Page 324; 608pp; English.

XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X<sub>1</sub>)-a-F<sub>1</sub>-(X<sub>2</sub>)-b, where: F<sub>1</sub> = an Fc domain; X<sub>1</sub> and X<sub>2</sub> = are each  
 CC independently selected from -(L<sub>1</sub>)-c-P<sub>1</sub>-(L<sub>2</sub>)-d-P<sub>2</sub>-(L<sub>1</sub>)-c-P<sub>1</sub>-  
 CC (L<sub>2</sub>)-d-P<sub>2</sub>-(L<sub>3</sub>)-e-P<sub>3</sub>- or -(L<sub>1</sub>)-c-P<sub>1</sub>-(L<sub>2</sub>)-d-P<sub>2</sub>-(L<sub>3</sub>)-e-P<sub>3</sub>-(L<sub>4</sub>)-f-P<sub>4</sub> where P<sub>1</sub>, P<sub>2</sub>,  
 CC P<sub>3</sub>, and P<sub>4</sub> = are each independently sequences of pharmacologically active  
 CC peptides; L<sub>1</sub>, L<sub>2</sub>, L<sub>3</sub>, and L<sub>4</sub> = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antineoplastic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAB5943 to AAB69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention

XX Sequence 36 AA;

Query Match 56.1%; Score 59.5; DB 3; Length 36;

Best Local Similarity 68.4%; Pred. No. 0.064; 1; Indels 5; Gaps 1;

Matches 13; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

3 IEGPTLRQWL-----HGNG 16  
 1 IEGPTLRQWLAAAGGNG 19

Db

RESULT 36

ABP51688 standard; peptide; 18 AA.

XX ABP51688;

XX 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:39.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 XX complementarity determining region; immunoglobulin; antianaemic;  
 XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
 XX Homo sapiens.  
 XX Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

XX 04-MAY-2001; 2001US-0288889P.

XX 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX WPI, 2002-566610/60.

XX N-PSDB; ABO73366.

XX A novel immunogen molecule comprising a region in which amino acid  
 XX residues corresponding to at least a portion of the complementary  
 XX determining region are replaced or fused with an erythropoietin or  
 XX thrombopoietin mimetic.

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antineoplastic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells; and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC AB073288 to AB073377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 55.7%; Score 59; DB 5; Length 18;  
 Best Local Similarity 91.7%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWL 12  
 | | | | | | | | | |  
 Db 1 LPIEGPTLRQWL 12

RESULT 37

ABP51677  
 ID ABP51677 standard; peptide; 18 AA.

XX ABP51677;

XX 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:61.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

XX complementarity determining region; immunoglobulin; antianaemic;

XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

XX Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

XX 04-MAY-2001; 2001US-0288889P.

XX 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdlish KS, Barbas-Frederickson S, Renshaw M;

XX WPI; 2002-566610/60.

XX A novel immunogen molecule comprising a region in which amino acid

XX residues corresponding to at least a portion of the complementary

XX determining region are replaced or fused with an erythropoietin or

XX thrombopoietin mimetic.

XX Claim 91; Page 57; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC AB073288 to AB073377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

XX Sequence 18 AA;

Query Match 55.7%; Score 59; DB 5; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 IEPTLRQWLHNGR 17  
 | | | | | | | | | |  
 Db 3 IEPTLRQWLARAR 17

RESULT 38

ABP51675  
 ID ABP51675 standard; peptide; 18 AA.

XX ABP51675;

XX 01-OCT-2002 (first entry)

XX TPO mimetic antibody related peptide graft SEQ ID NO:66.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

XX complementarity determining region; immunoglobulin; antianaemic;

XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

XX Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US047656.

XX 05-DEC-2000; 2000US-0251448P.

XX 04-MAY-2001; 2001US-0288889P.

XX 29-MAY-2001; 2001US-0294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdlish KS, Barbas-Frederickson S, Renshaw M;

XX WPI; 2002-566610/60.

XX A novel immunogen molecule comprising a region in which amino acid

XX residues corresponding to at least a portion of the complementary

XX determining region are replaced or fused with an erythropoietin or

XX thrombopoietin mimetic.

XX Example 4; Page 55; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (1) is contacted with haematopoietic  
 CC stem cells or their progenitors. (1) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention

SQ Sequence 18 AA;

Query Match 55.7%; Score 59; DB 5; Length 18;  
 Best Local Similarity 91.7%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWL 12  
 1 LPIEGPTLRQWL 12

RESULT 39

ADQ16619  
 ID ADQ16619 standard; peptide; 18 AA.

AC ADQ16619;

DT 09-SEP-2004 (first entry)

DE TPO mimetic peptide with random flanking residues SEQ ID NO:39.

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.

XX Unidentified.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;

DR WPI; 2004-460973/43.

DR N-PSDB; ADQ16620.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.

PS Example 1; SEQ ID NO 39; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

SQ Sequence 18 AA;

Query Match 55.7%; Score 59; DB 8; Length 18;

Best Local Similarity 91.7%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWL 12  
 1 LPIEGPTLRQWL 12

RESULT 40

ADQ16641  
 ID ADQ16641 standard; peptide; 18 AA.

AC ADQ16641;

DT 09-SEP-2004 (first entry)

DE TPO mimetic peptide with flanking amino acids SEQ ID NO:61.

KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KW immunotherapy; thrombocytopenia.

XX Unidentified.

PN WO2004050017-A2.

PD 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

PA (ALEX-) ALEXION PHARM INC.

PI Bowdish KS, Frederickson S, Renshaw M;

DR WPI; 2004-460973/43.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.

PS Example 6; SEQ ID NO 61; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide with flanking  
 CC residues.

PS Sequence 18 AA;

Query Match 55.7%; Score 59; DB 8; Length 18;  
 Best Local Similarity 73.3%; Pred. No. 0.035;  
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 IEGPTLRQWLHNGR 17  
 3 IEGPTLRQWLHNGR 17

RESULT 41

ADQ16646  
 ID ADQ16646 standard; peptide; 18 AA.

AC ADQ16646;

XX

DT	09-SEP-2004	(first entry)
XX	TPO mimetic peptide SEQ ID NO:65.	
DE		
KV	immunoglobulin; complementarity determining region; CDR; peptide mimetic;	
KW	erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;	
XX	immunotherapy; thrombocytopenia.	
XX		
OS	unidentified.	
XX		
PN	WO2004050017-A2.	
XX		
PD	17-JUN-2004.	
XX		
PF	17-NOV-2003; 2003WO-US036894.	
XX		
PR	02-DEC-2002; 2002US-00307724.	
XX		
PA	(ALEX-) ALEXION PHARM INC.	
XX		
PI	Bowditch KS, Frederickson S, Renshaw M,	
DR	MP1; 2004-460973/43.	
XX	N-PSDB; ADQ16645.	
PT	New immunoglobulin molecule comprising a region, where two	
XX	complementarity determining regions (CDRs) are replaced with EPO mimetic	
PT	or a TPO mimetic, useful for treating thrombocytopenia.	
XX		
PS	Example 4; SEQ ID NO 66; 107pp; English.	
XX		
CC	The invention relates to a novel immunoglobulin molecule or its fragment	
CC	comprising a region where amino acid residues corresponding to at least a	
CC	portion of a two complementarity determining regions (CDRs) are replaced	
CC	with a peptide mimetic selected from an erythropoietin (EPO) mimetic and	
CC	a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the	
CC	invention has immunosuppressive activity, and may have a use in	
CC	immunotherapy. The immunoglobulin molecule is useful for diagnosing or	
CC	treating thrombocytopenia as a result of chemotherapy, bone marrow	
CC	transplantation, or chronic diseases such as idiopathic thrombocytopenia.	
CC	The present sequence represents a TPO mimetic peptide of the invention.	
XX		
SO	Sequence 18 AA;	
Query Match	55.7%; Score 59; DB 8; Length 18;	
Best Local Similarity	91.7%; Pred. No. 0.035;	
Matches	11; Conservative 0; Mismatches 1; Indels 0; Gaps 0.	
OY	1 LAIEGPTLRQWL 12	
Db	1 LPIEGPTLRQWL 12	
RESULT 42		
ID	ADNS59819 standard; peptide; 22 AA.	
XX		
AC	ADNS59819;	
DT	01-JUL-2004 (first entry)	
XX		
DE	TMP peptide TWPI.	
XX		
KM	Haemostatic; antianaemic; immunosuppressive; platelet;	
KW	transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;	
KW	TMPI; c-mpl receptor; platelet precursor; megakaryocyte;	
KW	thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;	
KW	autoimmune haemolytic anaemia; Hughes' syndrome;	
XX	lupoid thrombocytopenia; linker.	
OS	Homo sapiens.	
XX		
PN	WO2003031589-A2.	

PD	17-APR-2003.
XX	
XX	11-OCT-2002; 2002MO-US032552.
PF	
XX	
XX	11-OCT-2001; 2001US-0328666P.
PR	
XX	10-OCT-2002; 2002US-00269806.
XX	
PA	(AMGE-) AMGEN INC.
XX	
PI	Min H, Sitney KC, Hartley C;
XX	
DR	WPI; 2003-403101/38.
XX	
PT	Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
XX	which stimulate the production of platelets and/or the production of
XX	platelet precursors, useful for treating thrombocytopenia.
XX	
PS	Example 6; Page 83; 126pp; English.
XX	
XX	The invention relates to a thrombopoietin mimetic peptide (TMP) (i) that
CC	binds to the c-mpl (mpl) receptor, and which stimulates the production of
CC	platelets and/or the production of platelet precursors, is new. Further
CC	disclosed is a composition of matter (ii) that binds to an mpl receptor,
CC	and a pharmaceutical composition comprising (iii) and a carrier. The
CC	pharmaceutical composition of the invention is useful for treating
CC	thrombocytopenia in an animal, and for increasing megakaryocytes or
CC	platelets in a patient. The TMP of the invention is useful for treating
CC	conditions involving a megakaryocyte and/or platelet deficiency, e.g.
CC	disease conditions involving thrombocytopenia such as aplastic anaemia,
CC	autoimmune thrombocytopenia, drug induced immune thrombocytopenia,
CC	autoimmune hemolytic anaemia, Hughes's syndrome and lupoid
CC	thrombocytopenia. The TMP of the invention is also useful for
CC	maintaining the viability or storage life of platelets and/or
CC	megakaryocytes and its derived cells. The compounds demonstrate an
CC	improved ability to bind to and/or trigger transmembrane signal through,
CC	i.e. activating the mpl receptor the compounds have superior
CC	thrombopoietic activity, i.e. the ability to stimulate, in vivo and in
CC	vitro, the production of platelets and/or megakaryocytopenic activity,
CC	i.e. the ability to stimulate, in vivo and in vitro, the production of
CC	platelet precursors. Further, certain of the compounds also exhibit
CC	superior therapeutic properties, such as improved plasma half-life,
CC	biological activity and in vivo circulation time. The current sequence
CC	represents a TMP peptide of the invention to which a two amino acid "cap"
CC	has been added to the carboxy terminal to increase peptide affinity.
XX	
XX	
SQ	Sequence 22 AA;
XX	
XX	
Query Match	55.7%; Score 59; DB 7; Length 22;
Best Local Similarity	64.7%; Pred. NO. 0.044;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;	
QY	3 IEGPLRLQWLHGNGRDT 19
Db	5 IEGPLRLQWLAAALLET 21
XX	
XX	
RESULT 43	
AD016705	
ID	AD016705 standard; protein; 128 AA.
XX	
AC	AD016705;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Modified immunoglobulin clone 116 HC variable region SEQ ID NO:125.
XX	
KM	immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW	erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KX	immunotherapy; thrombocytopenia.
XX	
OS	Synthetic.
XX	

PT	New immunoglobulin molecule comprising a region, where two complementary determining regions (CDRs) are replaced with EPO mimetic or a TPO mimetic, useful for treating thrombocytopenia.
XX	
PS	Example 8; SEQ ID NO 124; 107bp; English.
XX	
CC	The invention relates to a novel immunoglobulin molecule or its fragment comprising a region where amino acid residues corresponding to at least a portion of a two complementarity determining regions (CDRs) are replaced with a peptide mimetic selected from an erythropoietin (EPO) mimetic and a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the invention has immunosuppressive activity, and may have a use in immunotherapy. The immunoglobulin molecule is useful for diagnosing or treating thrombocytopenia as a result of chemotherapy, bone marrow transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC	The present sequence represents immunoglobulin clone 116 heavy chain.
CC	
XX	
SO	Sequence 225 AA;
Query Match	55.7%; Score 59; DB 8; Length 225;
Best Local Similarity	91.7%; Pred. No. 0.59; 1; Indels 0; Gaps 0;
Matches	11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy	1 LAIEGPTLRQWL 12                 100 LPIEGPTLRQWL 111
Ds	
RESULT 45	
ID	ABP51695 standard; protein; 472 AA.
XX	ABP51695
AC	ABP51695;
XX	
DT	01-OCT-2002 (first entry)
XX	
DE	5G1.1-TPO heavy chain amino acid sequence SEQ ID NO:67.
XX	
KW	TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region; complementarity determining region; immunoglobulin; antianaemic; haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200246238-A2.
XX	
PD	13-JUN-2002.
XX	
PF	05-DEC-2001; 2001WO-US047656.
XX	
PR	05-DEC-2000; 2000US-0251448P.
PR	04-MAY-2001; 2001US-0288889P.
PR	29-MAY-2001; 2001US-0294068P.
XX	
PA	(ALEX-) ALEXION PHARM INC.
XX	
PI	Bowditch KS, Barbas-Frederickson S, Renshaw M;
XX	
DR	WPI; 2002-566610/60.
XX	
DR	N-P8DB; ABQ73374.
XX	
XX	
PT	A novel immunogen molecule comprising a region in which amino acid residues corresponding to at least a portion of the complementary determining region are replaced with an erythropoietin or thrombopoietin mimetic.
XX	
PS	Example 4; Fig 13A; 113bp; English.
XX	
CC	The present invention describes an immunoglobulin molecule or its fragment (I) comprising a region where amino acid residues corresponding to at least a portion of the complementary determining region (CDR) are replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC anti-nausea, haemostatic and nephroprotective activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease.  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 XX

Sequence 472 AA;

Query Match 55.7%; Score 59; DB 5; Length 472;  
 Best Local Similarity 91.7%; Pred. No. 1.3;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWL 12  
 | |||||  
 Db 118 LPIEGPTLRQWL 129

Search completed: September 1, 2005, 16:12:13  
 Job time : 88.3453 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:57:33 ; Search time 14.4892 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-11

Perfect score: 106  
Sequence: 1 LAIEGPTLRQMLHNGRDT 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:1\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	# Query Match	Length	DB ID	Description
1	50	47.2	691	2 A54741	erythrocyte membra
2	50	47.2	721	2 A39707	erythrocyte membra
3	50	47.2	973	2 AB2340	hypothetical prote
4	50	47.2	1774	2 S13178	6-methylsalicylic
5	49	46.2	419	1 ESECRM	erythromycin ester
6	47	44.3	349	2 B87251	molybdenum cofacto
7	47	44.3	391	2 E83151	hypothetical prote
8	46.5	43.9	602	2 T45278	oligopeptide ABC t
9	46	43.4	434	2 T31313	glutamate-1-semial
10	45.5	42.9	333	2 A36925	transcription acti
11	45	42.5	235	2 C83822	hypothetical prote
12	45	42.5	302	2 JN0143	catechol 1,2-dioxy
13	45	42.5	304	2 C90453	hypothetical prote
14	45	42.5	492	2 T01086	probable serine/th
15	44	41.5	209	2 A13455	transcription regu
16	44	41.5	229	2 JC7219	nuclear protein SR
17	44	41.5	278	2 F84127	hypothetical prote
18	44	41.5	331	2 B48445	glyceraldhyde-3-P
19	44	41.5	344	2 AE3379	molybdenum cofacto
20	44	41.5	496	2 S25091	cruciferin Bnc2 -
21	44	41.5	571	2 A10506	probable sulfatase
22	44	41.5	972	2 T49773	related to actin-i
23	44	41.5	1712	2 CGH028	collagen alpha 2(I
24	44	41.5	3430	1 GNMWV	genome polyprotein
25	44	41.5	3433	1 GNMWV	genome polyprotein
26	43.5	41.0	296	2 AG0147	probable membrane
27	43	40.6	218	2 H82539	protein-L-leasapar
28	43	40.6	313	2 A45822	beta-lactamase (BC
29	43	40.6	352	2 B69901	fatty-acid desatur

30	43	40.6	481	2 T05270	probable serine/th
31	43	40.6	664	2 G89894	protein kinase [im
32	43	40.6	841	2 A43254	protein-tyrosine-P
33	43	40.6	1023	2 E71376	conserved hypothet
34	42	39.6	104	2 B82797	conserved hypothet
35	42	39.6	106	2 AC3086	sarcosine oxidase
36	42	39.6	106	2 F98200	sarcosine oxidase
37	42	39.6	218	2 D83161	hypothetical prote
38	42	39.6	273	2 T44657	protein GP80 (limp
39	42	39.6	335	2 B72053	glyceraldhyde 3-P
40	42	39.6	335	2 B86568	glyceraldhyde 3-P
41	42	39.6	335	2 S43339	glyceraldhyde 3-P
42	42	39.6	339	2 A30754	hypothetical prote
43	42	39.6	407	2 A86298	hypothetical prote
44	42	39.6	422	2 P66826	hypothetical prote
45	42	39.6	433	2 S51837	glyceraldhyde-3-P
46	42	39.6	433	2 S51836	glyceraldhyde-3-P
47	42	39.6	457	2 D70901	probable tmu prote
48	42	39.6	473	2 E84853	hypothetical prote
49	42	39.6	926	2 B84642	hypothetical prote
50	42	39.6	1019	2 T11560	pol polyprotein -
51	42	39.6	1022	2 T51257	calmodulin-binding
52	42	39.6	1022	2 T50928	calmodulin-binding
53	42	39.6	1075	2 D70568	hypothetical prote
54	42	39.6	1075	2 D70568	hypothetical prote
55	42	39.6	2357	2 A59249	class VII unconven
56	42	39.6	2843	1 RBHUP	adenomatous polyp
57	42	39.6	2845	1 I48505	adenomatous polyp
58	41.5	39.2	495	2 S00657	apoprotein(a) (EC
59	41.5	39.2	521	2 AG0103	hypothetical prote
60	41.5	39.2	521	2 H86298	hypothetical prote
61	41.5	39.2	732	2 T45429	polyposphate kina
62	41.5	39.2	742	2 E70673	probable ppk prote
63	41	38.7	306	2 P97120	translation elonga
64	41	38.7	349	2 B97912	tdpglucose 4,6-de
65	41	38.7	417	2 S07286	hypothetical prote
66	41	38.7	428	2 F85849	probable integrase
67	41	38.7	428	2 E91005	probable integrase
68	41	38.7	450	2 T01711	probable serine/th
69	41	38.7	497	2 A86146	hypothetical prote
70	41	38.7	518	2 AD2315	hypothetical prote
71	41	38.7	525	2 C69794	glutamate synthase
72	41	38.7	549	2 G97614	ABC transporter (A
73	41	38.7	549	2 AE2837	hypothetical prote
74	41	38.7	555	2 S56946	probable membrane
75	41	38.7	584	2 AC3321	ABC transporter AT
76	41	38.7	600	2 T00759	hypothetical prote
77	41	38.7	631	2 B87250	dnak protein (limp
78	41	38.7	703	2 AC2430	hypothetical prote
79	41	38.7	905	2 T02205	lu-ECAM-1 protein
80	41	38.7	1149	2 T30869	adenosinetriphosph
81	41	38.7	1149	2 T30869	probable adenosine
82	41	38.7	1452	1 S17669	protein-tyrosine-P
83	41	38.7	1452	1 S17670	protein-tyrosine-P
84	41	38.7	1478	2 C82689	helicase, ATP depe
85	41	38.7	1639	2 T14181	peptide synthetase
86	41	38.7	2569	2 T14164	peptide synthetase
87	40.5	38.2	98	2 A70301	ribosomal protein
88	40.5	38.2	110	2 G90584	50S ribosomal prot
89	40.5	38.2	124	2 D71355	probable ribosomal
90	40.5	38.2	345	2 H71358	conserved hypothet
91	40.5	38.2	447	2 G95068	cysteiny1-CRNA syn
92	40.5	38.2	469	2 T48458	cysteine-tRNA liga
93	40.5	38.2	1175	2 T46124	8-amino-7-oxononan
94	40.5	38.2	1387	2 A96771	hypothetical prote
95	40.5	38.2	1736	2 T00391	hypothetical prote
96	40	37.7	93	2 E70967	phosphothricin N
97	40	37.7	154	2 AE3299	hypothetical prote
98	40	37.7	175	2 F69745	hypothetical prote
99	40	37.7	211	2 B95041	hypothetical prote
100	40	37.7	215	2 E96533	hypothetical prote

## ALIGNMENTS

## RESULT 1

A54741

erythrocyte membrane band 4.2 protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004

C/Accession: A54741.148901

R/Korsgren, C.; Cohen, C.M.

Genomics 21, 478-485, 1994

A/Title: cDNA sequence, gene sequence, and properties of murine pallidin (band 4.2), the

A/Reference number: A54741; MUID:95048323; PMID:7959722

A/Accession: A54741

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-691 &lt;KOR&gt;

A/Cross-references: UNIPROT:P49222; GB:U04055

A/Note: authors translated the codon TAC for residue 129 as Ile, and GCT for residue 352

R./Rybicki, A.C.; Schwartz, R.S.; Qiu, J.J.; Gilman, J.G.

Mamm. Genome 5, 438-445, 1994

A/Title: Molecular cloning of mouse erythrocyte protein 4.2: a membrane protein with str

A/Reference number: I48901; MUID:95003352; PMID:7919657

A/Accession: I48901

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-352, 'A', 354-620, 'S', 622-691 &lt;RES&gt;

A/Cross-references: EMBL:U03487; NID:g424119; PIDN:AAA62275.1; PID:g424120

C/Superfamily: protein-glutamine gamma-glutamyltransferase

C/Keywords: blocked amino end; lipoprotein; myristylation

F./2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

Query Match 47.2%; Score 50; DB 2; Length 691;

Best Local Similarity 75.0%; Pred. No. 9.1;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 6 PTLROWLHGNGR 17

250 PTLROWLHGNGR 261

## RESULT 2

A39707

erythrocyte membrane band 4.2 protein - human

N/Alternate names: pallidin

C/Species: Homo sapiens (man)

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: A39707; A34865; B34865; A34883

R/Korsgren, C.; Cohen, C.M.

Proc. Natl. Acad. Sci. U.S.A. 88, 4840-4844, 1991

A/Title: Organization of the gene for human erythrocyte membrane protein 4.2: structural

A/Reference number: A39707; MUID:91271288; PMID:2052563

A/Accession: A39707

A/Molecule type: DNA

A/Residues: 1-721 &lt;KOR1&gt;

A/Cross-references: UNIPROT:P16452; GB:I06519; NID:g306738; PIDN:AAA52385.1; PID:g306740

A/Experimental source: cell type erythrocyte; tissue type peripheral blood; tissue lib h

R./Sung, L.A.; Chien, S.; Chang, L.S.; Lambert, K.; Bliss, S.A.; Bouhasnia, E.E.; Nagel,

Proc. Natl. Acad. Sci. U.S.A. 87, 955-959, 1990

A/Title: Molecular cloning of human protein 4.2: a major component of the erythrocyte me

A/Reference number: A34865; MUID:90138995; PMID:1689063

A/Accession: A34865

A/Molecule type: mRNA

A/Residues: 1-364, 'KRGGLPC', 371-379, 'H', 381-405, 'L', 407-721 &lt;SUN1&gt;

A/Cross-references: GB:M30647; NID:g189433; PIDN:AAA36401.1; PID:g189434

A/Accession: B34865

A/Molecule type: mRNA

A/Residues: 1-33, 34-364, 'KRGGLPC', 371-379, 'H', 381-405, 'L', 407-721 &lt;SUN2&gt;

A/Cross-references: GB:M30646; NID:g189435; PIDN:AAA36402.1; PID:g189436

A/Experimental source: isolate sickle cell patient; cell type reticulocyte

A/Note: parts of this sequence were determined by protein sequencing

R/Korsgren, C.; Lawler, J.; Lambert, S.; Speicher, D.; Cohen, C.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 613-617, 1990

A/Title: Complete amino acid sequence and homologies of human erythrocyte membrane prote

A/Reference number: A34883; MUID:90138879; PMID:2300550

A/Accession: A34883

A/Molecule type: mRNA

A/Residues: 1-3, 34-721 &lt;KOR2&gt;

A/Cross-references: GB:M29399; NID:g182083; PIDN:AAA35798.1; PID:g182084

C/Comment: This protein is a major constituent of the erythrocyte membrane. It apparentl

C/Genetics:

A/Gene: GDB:EPB42, PA

A/Cross-references: GDB:127385; OMIM:177070

A/Map position: 15q15-15q15

C/Superfamily: protein-glutamine gamma-glutamyltransferase

C/Keywords: alternative splicing; blocked amino end; glycoprotein; lipoprotein; myristyl

F./2/2/Product: erythrocyte membrane band 4.2 protein, long splice form #status predict

F./2-3/34-721/Product: erythrocyte membrane band 4.2 protein, short splice form #status p

F./298-316/Domain: transmembrane #status predicted &lt;IRM&gt;

F./518-520/Region: cell attachment (R-G-D) motif

F./2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F./103,420,447,529,604,705/Binding site: carbohydrate (Asn) (covalent) #status predicted

F./278/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predic

Query Match 47.2%; Score 50; DB 2; Length 721;

Best Local Similarity 75.0%; Pred. No. 9.5;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 6 PTLROWLHGNGR 17

280 PTLROWLHGNGR 291

## RESULT 3

AB2340

hypothetical protein alr4273 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C/Accession: AB2340

R./Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A/Title: Complete genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AB2340

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-973 &lt;KUN&gt;

A/Cross-references: UNIPROT:O8YPC5; GB:BA000019; PIDN:BA875972.1; PID:g17133408; GSPDB:G

A/Experimental source: strain PCC 7120

C/Genetics:

A/Gene: alr4273

Query Match 47.2%; Score 50; DB 2; Length 973;

Best Local Similarity 72.7%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 6 PTLROWLHGNG 16

649 PTLROWLHGNG 659

## RESULT 4

S13178

6-methylsalicylic acid synthase - Penicillium griseofulvum

C/Species: Penicillium griseofulvum

C/Date: 21-Nov-1993 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004

C/Accession: S13178

R./Beck, J.; Ripka, S.; Siegner, A.; Schilz, E.; Schweitzer, E.

Eur. J. Biochem. 192, 487-498, 1990

A/Title: The multifunctional 6-methylsalicylic acid synthase gene of Penicillium patulum

A/Reference number: S13178; MUID:91006137; PMID:2209605

A/Accession: S13178

A/Status: preliminary



A: Molecule type DNA  
A: Residues: 1-1774 <BRC>  
A: Cross-references: UNIPROT: P22367; GB: X5776; NID: g3211; PID: CA13295.1; PID: g3212  
C: Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoad  
homology; [acyl-carrier-protein] 5-malonyltransferase homology  
C: Keywords: carrier protein  
E: 54-455/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
E: 567-844/Domain: [acyl-carrier-protein] 5-malonyltransferase homology <AMT>  
E: 1412-1605/Domain: short-chain alcohol dehydrogenase homology <SAD2>  
E: 1658-1768/Domain: acyl-chain alcohol protein homology <ACP2>

Query Match	47.2%;	Score 50;	DB 2;	Length 1774;
Best Local Similarity	44.4%;	Pred. No. 25;		
Matches	8;	Conservative	4;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
QY      1 LAIEGPTLRQWLHNGRD 18
         ||: ||| |: ||
Db      488 LALQAKTLRDWMTAEGKD 505
```

RESULT 5  
ESEC RM

A:Gene: ereB  
C:Function:  
A:Description: erythromycin esterase [validated, MUID:86259072]  
C:Superfamily: erythromycin esterase, type II  
C:Keywords: antibiotic resistance; carboxylic ester hydrolase

Query Match	46.2%	Score 49	DB 1	Length 419
Best Local Similarity	46.7%	Pred. No. 7.5		
Matches	7	Conservative	2	Mismatches 6
				Indels 0
				Gaps 0
OY	4	EGPTLRQWLKNGNRD	18	
		:		
		:		
Db	79	EGQIINNMIHQGSTD	93	

RESULT 6  
BB7251  
molybdenum cofactor biosynthesis protein A [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 16-Aug-2004

**A1:Gene:** CC0018  
**C1:Superfamily:** Molybdenum cofactor Molybdenum cofactor precursor Z biosynthesis protein  
**Query Match**      **44.3%**;   **Score 477**;   **DB 2**;   **Length 349**;

Best Local Similarity 60.0%; Pred. No. 13;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	4	E	G	P	T	R	Q	W	L	H	G	N	R	D	18
Db	191	E	I	P	A	L	I	Q	W	A	H	G	R	G	205

RESULT 7  
E83151  
hypochlorite protein PA3949 [imported] - Pseudomonas aeruginosa (strain PA01)  
C1:Species: Pseudomonas aeruginosa  
C2:Accession: E83151  
C3:Date: 15 Oct 2006  
C4:Host: shrimp  
C5:Lab: D0  
C6:Proj: T1  
C7:2004

A;Cross-references: UNIPROT:Q9HX7; GB:AE0040813; GB:AE004091; NID:g95950134; PION:AA60733  
A;Experimental source: strain PAO1  
C;Genetics:  
A;Gene: PA3949

```
QY      4  EGPTRLRQWLHGNRDT 19
          : |||: ||| ||
Db      64 DADALRAWIHGLGIDT 79
```

RESULT 8  
T45278  
oligopeptide ABC transport protein bldKB [imported] - Streptomyces coelicolor  
CjSpecies: Streptomyces coelicolor  
CjDate: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004  
CjAccession: T45278  
R.Nodwell J., R.; McGovern, K.; Losick, R.  
submitted to the EMBL Data Library, August 1996  
A>Description: An oligopeptide permease responsible for the import of an extracellular  
A.Reference number: Z22954

A;Note: bldKB  
C;Function:  
A;Description: involved in aerial mycelium formation  
C;Keywords: oligopeptide transport

QY 4 EGPT-LRQWLHGNG 16  
:|||:||||:  
Db 174 DGPTYLQQLSGDG 187

RESULT 9  
T1313  
glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) - Cenarchaeum symbiosum  
C1Species: Cenarchaeum symbiosum

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
 C/Accession: T31313  
 R/Schleper, C.; Delong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.  
 J. Bacteriol. 180, 5003-5009, 1998  
 A>Title: Genomic analysis reveals chromosomal variation in natural populations of the *un*  
 A/Reference number: Z20994; MUID:98422450; PMID:9748430  
 A/Accession: T31313  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-434 <SCH>  
 A/Cross-references: UNIPROT:O74061; EMBL:AF083072; NID:g3599393; PID:g3599399; PIND:AA6  
 C/Genetics:  
 A/Note: gsat  
 C/Function:  
 A/Description: heme biosynthesis  
 C/Superfamily: ornithine-oxo-acid aminotransferase  
 C/Keywords: intramolecular transferase; isomerase

Query Match 43.4%; Score 46; DB 2; Length 434;  
 Best Local Similarity 41.2%; Pred. No. 23;  
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIEPTLRQWLGNGRD 18  
 :||| :||| :  
 DB 81 AVEGQLRGWIGTANE 97

RESULT 10  
 A36925  
 transcription activator LysR-type Cbdr - Xanthobacter flavus  
 C/Species: Xanthobacter flavus  
 C/Date: 04-Nov-1994 #sequence\_revision 04-Nov-1994 #text\_change 09-Jul-2004  
 C/Accession: A36925; S13578; S35408  
 R/Van den Bergh, E.R.E.; Dijkhuizen, L.; Meijer, W.G.  
 J. Bacteriol. 175, 6097-6104, 1993  
 A>Title: Cbdr, a LysR-type transcriptional activator, is required for expression of the  
 A/Reference number: A36925; MUID:94012468; PMID:8407781  
 A/Accession: A36925  
 A>Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-333 <VAN>  
 A/Cross-references: UNIPROT:P25545; EMBL:Z22705; NID:g297851; PIND:CAA80406.1; PID:g5618  
 R/Meijer, W.G.; Arberg, A.C.; Enequist, H.G.; Terpstra, P.; Lidstrom, M.E.; Dijkhuizen,  
 M.J. Gen. Genet. 225, 320-330, 1991  
 A>Title: Identification and organization of carbon dioxide fixation genes in Xanthobacte  
 A/Reference number: S13573; MUID:91172133; PMID:1500916  
 A/Accession: S13578  
 A/Molecule type: DNA  
 A/Residues: 1-150 <MEI>  
 A/Cross-references: EMBL:X17252  
 C/Genetics:  
 A/Gene: cbdr  
 A/Start codon: GTG  
 C/Superfamily: transcription activator LysR-type  
 C/Keywords: DNA binding; transcription regulation

Query Match 42.9%; Score 45.5; DB 2; Length 333;  
 Best Local Similarity 52.6%; Pred. No. 20;  
 Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 LATEG-PTLRQWLGNGRD 18  
 :||| :||| :|||  
 DB 262 LEVEGLPVRQWLAVARD 280

RESULT 11  
 C83822  
 hypothetical protein BH1379 [imported] - Bacillus halodurans (strain C-125)  
 C/Species: Bacillus halodurans  
 C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C/Accession: C83822  
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A/Reference number: A83650; MUID:20512582; PMID:11058132  
 A/Accession: C83822  
 A>Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-235 <STO>  
 A/Cross-references: UNIPROT:Q9KD40; GB:AP001511; GB:BA000004; NID:g10173727; PIND:BA050  
 A/Experimental source: strain C-125  
 C/Genetics:  
 A/Gene: BH1379

Query Match 42.5%; Score 45; DB 2; Length 235;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 IEPTLRQWLGNG 16  
 :||| :||| :|||  
 DB 125 VHGKAIQWLSNDG 138

RESULT 12  
 JN0143  
 catechol 1,2-dioxygenase (EC 1.13.11.1) - Pseudomonas sp. plasmid EST1001  
 C/Species: Pseudomonas sp.  
 C/Date: 05-Mar-1993 #sequence\_revision 05-Mar-1993 #text\_change 09-Jul-2004  
 C/Accession: JN0143  
 R/Kivisaar, M.; Kasaak, L.; Nurk, A.  
 Gene 98, 15-20, 1991  
 A>Title: Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, pheB,  
 A/Reference number: JN0143; MUID:91192610; PMID:2013408  
 A/Accession: JN0143  
 A/Molecule type: DNA  
 A/Residues: 1-302 <KIV>  
 A/Cross-references: UNIPROT:P31019; GB:M57500; NID:g145127; PIND:AA64900.1; PID:g145129  
 C/Genetics:  
 A/Gene: pheB  
 A/Genome: plasmid  
 C/Superfamily: catechol 1,2-dioxygenase  
 C/Keywords: aromatic hydrocarbon catabolism; oxidoreductase

Query Match 42.5%; Score 45; DB 2; Length 302;  
 Best Local Similarity 56.2%; Pred. No. 22;  
 Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 4 EGPTLRQWLGNGRDT 19  
 :||| :||| :|||  
 DB 130 DGETW--MLHGQVRDT 143

RESULT 13  
 C90453  
 hypothetical protein hpcE-2 [imported] - Sulfolobus solfataricus  
 C/Species: Sulfolobus solfataricus  
 C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C/Accession: C90453  
 R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-  
 jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P  
 submitted to Genbank, April 2001  
 A/Description: Sulfolobus solfataricus complete genome.  
 A/Reference number: A99139  
 A/Accession: C90453  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-304 <KUN>  
 A/Cross-references: UNIPROT:Q97V63; GB:AE006641; NID:g13816109; PIND:AAK42882.1; GSPDB:G  
 C/Genetics:  
 A/Gene: hpcE-2

Query Match 42.5%; Score 45; DB 2; Length 304;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGGPTLRQWLHGNGRDT 19  
DB 187 EMPYGRWVHGKMDT 202

## RESULT 14

T01086  
Probable serine/threonine-specific protein kinase (EC 2.7.1.-) T10P11.10 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #ext\_change 09-Jul-2004  
C:Accession: T01086  
R:Kaplan, N.; Johnson, D.; Schütz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.; Martensen, R.; Chen, E.Y.; Wilson, R.; McComble, W.R.  
Submitted to the EMBL Data Library, November 1998  
A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.  
A:Reference number: Z14248  
A:Accession: T01086  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-492 <KAP>  
A:Cross-references: UNIPROT:Q22764; EMBL:AC002330; NID:g2262135; PID:g2262143  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Note: T10P11.10  
C:Keywords: phosphotransferase

Query Match 42.5%; Score 45; DB 2; Length 492;  
Best Local Similarity 42.9%; Pred. No. 37;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 IEGETLRQWLHGNG 16  
DB 238 VDNQNLQWIMHGCG 251

## RESULT 15

A13455  
transcription regulator, tetr family BMEI1631 [imported] - Brucella melitensis (strain 1  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #ext\_change 09-Jul-2004  
C:Accession: A13455  
R:Delvecchio, V.G.; Kapactral, V.; Redkar, R.J.; Petra, G.; Mujter, C.; Loe, T.; Ivanova, R.; Masur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letenski, J.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: A13455  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-209 <KUR>  
A:Cross-references: UNIPROT:Q8YF91; GB:AB008917; PIDN:AL52812.1; PID:g17983650; GSPDB:G  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI1631  
A:Map position: 1  
C:Superfamily: probable transcription repressor mtrr

Query Match 41.5%; Score 44; DB 2; Length 209;  
Best Local Similarity 57.1%; Pred. No. 21;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHG 14  
DB 167 LANVGGRFLNEMWLG 180

## RESULT 16

JC7219  
nuclear protein SR-25 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #ext\_change 09-Jul-2004  
C:Accession: JC7219

R:Sasahara, K.; Yamaoka, T.; Moritani, M.; Tanaka, M.; Iwahana, H.; Yoshimoto, K.; Miyag  
Biochem. Biophys. Res. Commun. 269, 444-450, 2000  
A:Title: Molecular cloning and expression analysis of a putative nuclear protein, SR-25  
A:Reference number: JC7219; MUID:20175222; PMID:10708573  
A:Accession: JC7219  
A:Molecule type: mRNA  
A:Residues: 1-229 <SAS>

A:Cross-references: UNIPROT:Q9JW93; DBJ:AB035383; NID:g7619895; PIDN:BA094743.1; PID:g7  
A:Experimental source: M106 cell line  
C:Comment: This protein is a highly hydrophilic nuclear protein with a serine-arginine  
A:splicing factors.  
C:Keywords: nucleus; RNA processing

Query Match 41.5%; Score 44; DB 2; Length 229;  
Best Local Similarity 47.1%; Pred. No. 23;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIEGPTLRQWLHGNGRD 18  
DB 128 ALRPSLDQMHSAGED 144

## RESULT 17

F84127  
hypothetical protein BH3822 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #ext\_change 09-Jul-2004  
C:Accession: F84127  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: AB3650; MUID:20512582; PMID:11051132  
A:Accession: F84127  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <STO>  
A:Cross-references: UNIPROT:Q9K6A7; GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA07  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3822

Query Match 41.5%; Score 44; DB 2; Length 278;  
Best Local Similarity 70.0%; Pred. No. 29;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 PTLRQWLHNGN 15  
DB 48 PTLRAVWHN 57

## RESULT 18

B48445  
glyceraldhyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania m  
C:Species: Leishmania mexicana  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #ext\_change 09-Jul-2004  
C:Accession: B48445; S25142  
R:Hannaeert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendes, F.R.; Michels, P.A.M.  
Mol. Biochem. Parasitol. 55, 115-126, 1992  
A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate  
A:Reference number: A48445; MUID:93063042; PMID:1435864  
A:Accession: B48445  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-331 <HAN>  
A:Cross-references: UNIPROT:Q01558; EMBL:X65220; NID:g95552; PIDN:CAA46323.1; PID:g95553  
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase  
C:Keywords: oxidoreductase

Query Match 41.5%; Score 44; DB 2; Length 331;  
Best Local Similarity 42.9%; Pred. No. 35;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 IEGETLRQWLHGNG 16

Db 185 VDPSTKDMRGRC 198

## RESULT 19

AE3379  
molybdenum cofactor biosynthesis protein A [imported] - Brucella melitensis (strain 16M)  
C/Species: Brucella melitensis  
C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 16-Aug-2004  
C/Accession: AE3379  
R/DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, R.; Selkov, E.; Elzer, P.H.; Haglue, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A/Reference number: AD3252; PMID:11756688  
A/Accession: AE3379  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-344 <KUR>  
A/Cross-references: UNIPROT:Q8YGY6; GB:AE008917; PIDN:AA152200.1; PID:gl7982982; GSPDB:G  
A/Experimental source: strain 16M  
C/Genetics:  
A/Gene: BME11019  
A/Map position: 1  
C/Superfamily: Molybdenum cofactor Molybdenum cofactor precursor Z biosynthesis protein

Query Match 41.5%; Score 44; DB 2; Length 344;  
Best Local Similarity 53.3%; Pred. No. 36;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGPTLRQWLHGNGRD 18  
Db 184 EIEPLTRMAGRGMD 198

## RESULT 20

S25091  
cruciferin Bnc2 - rape  
C/Species: Brassica napus (rape)  
C/Date: 04-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 09-Jul-2004  
C/Accession: S25091  
R/Breen, J.P.; Crouch, M.L.  
Plant Mol. Biol. 19, 1049-1055, 1992  
A/Title: Molecular analysis of a cruciferin storage protein gene family of Brassica napus  
A/Reference number: S25090; MUID:92379259; PMID:151129  
A/Accession: S25091  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-496 <BRE>  
A/Cross-references: UNIPROT:P33524; EMBL:X59295; NID:gl7791; PIDN:CAA41985.1; PID:g76292  
C/Genetics:  
A/Gene: Bnc2  
A/Intons: 95/1; 222/2; 362/3  
C/Superfamily: glycinin  
C/Keywords: seed; storage protein

Query Match 41.5%; Score 44; DB 2; Length 496;  
Best Local Similarity 47.1%; Pred. No. 54;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LATEGPTLRQWLHGNGR 17  
Db 226 LAGKNPOGOSWHLGRQ 242

## RESULT 21

AI0506  
probable sulfatase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strai  
C/Species: Salmonella enterica subsp. enterica serovar Typhi  
A/Note: this species has also been called Salmonella typhi  
C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C/Accession: AI0506  
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,  
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A/Reference number: AB0502; MUID:21534947; PMID:11677608  
A/Accession: AI0506  
A/Status: preliminary  
A/Molecule type: DNA

A/Residues: 1-571 <PAR>  
A/Cross-references: GB:AL513382; PIDN:CAD01193.1; PID:gl6501322; GSPDB:GN00176  
C/Genetics:  
A/Gene: STY0046

Query Match 41.5%; Score 44; DB 2; Length 571;  
Best Local Similarity 58.3%; Pred. No. 63;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 EGPTLRQWLHGNG 15  
Db 308 EDPYKDWLHIN 319

## RESULT 22

T49773  
related to actin-interacting protein AIP3 [imported] - Neurospora crassa  
N/Alternate names: protein B9J10.100  
C/Species: Neurospora crassa  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: T49773  
R/Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25022  
A/Accession: T49773  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-972 <SCH>  
A/Cross-references: UNIPROT:Q9P571; EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.100  
A/Experimental source: BAC clone B9J10; strain OR74A  
C/Genetics:  
A/Gene: NCSP:B9J10.100  
A/Map position: 6  
A/Intons: 29/3; 161/1; 329/1

Query Match 41.5%; Score 44; DB 2; Length 972;  
Best Local Similarity 61.5%; Pred. No. 1.1e+02;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 TLQWLHGNGRDT 19  
Db 59 TLTQWSRGVATDT 71

## RESULT 23

CGH028  
collagen alpha 2(IV) chain precursor - human  
N/Alternate names: procollagen alpha 2(IV) chain  
C/Species: Homo sapiens (man)  
C/Date: 07-Jun-1990 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
C/Accession: A32024; S00007; S02624; S00246; S17678; S16911; B32117; S16877; S00165; S39  
R/Hoslika, S.L.; Tryggvason, K.  
J. Biol. Chem. 263, 19488-19493, 1988  
A/Title: The complete primary structure of the alpha2 chain of human type IV collagen an  
A/Reference number: A32024; MUID:89066769; PMID:3198637  
A/Accession: A32024  
A/Molecule type: mRNA  
A/Residues: 1-1712 <HOS1>  
A/Cross-references: UNIPROT:P08572; EMBL:U04210; EMBL:X05610; GB:W20753; NID:g29550; PID  
R/Hoslika, S.L.; Kurkinen, M.; Tryggvason, K.  
FEBS Lett. 216, 281-286, 1987  
A/Title: Nucleotide sequence coding for the human type IV collagen alpha-2 chain cDNA re  
ated region.  
A/Reference number: S00007; MUID:87219158; PMID:3582677

A:Accession: S00007  
 A:Molecule type: mRNA  
 A:Residues: 1254-1398, 'V', 1400-1712 <HOS2>  
 A:Cross-references: EMBL:J04210; EMBL:X05610; GB:M20753; NID:G29550; PIDN:CAA29098.1; PI  
 A>Note: 1399-116 was also found  
 R:Hoeflika, S.L.; Trygvaason, K.  
 FEBS Lett. 224, 297-305, 1987  
 A:Title: Extensive structural differences between genes for the alpha(1) and alpha(2) ch  
 A:Reference number: S02624; MUID:88083553; PMID:2826228  
 A:Accession: S02624  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1347-1350, 1377-1383, 1426-1432, 1465-1471, 1529-1535, 1625-1630 <HOS3>  
 A>Note: complete nucleotide sequence not shown  
 R:Brazel, D.; Pollner, R.; Oberbauer, I.; Kuehn, K.  
 Eur. J. Biochem. 172, 35-42, 1988  
 A:Title: Human basement membrane collagen (type IV): the amino acid sequence of the alph  
 A:Reference number: S00246; MUID:88151998; PMID:3345760  
 A:Accession: S00246  
 A:Molecule type: mRNA  
 A:Residues: 1-682, 'G', 684-1043 <BRA>  
 A:Cross-references: EMBL:X05562; NID:G30075; PIDN:CAA29076.1; PID:G30076  
 R:Oberbauer, I.  
 submitted to the EMBL Data Library, June 1987  
 A:Reference number: S17678  
 A:Accession: S17678  
 A:Molecule type: mRNA  
 A:Residues: 1-470, 'P', 472-682, 'G', 684-1043 <OBR>  
 A:Cross-references: EMBL:X05562; NID:G30075; PIDN:CAA29076.1; PID:G30076  
 R:Poersch, E.; Pollner, R.; Kuehn, K.  
 EMBO J. 7, 2687-2695, 1988  
 A:Title: The genes for the alpha(IV) and alpha2(IV) chains of human basement membrane  
 A:Reference number: S02738; MUID:89030632; PMID:2846280  
 A:Accession: S16911  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-33 <POE>  
 A:Cross-references: EMBL:X12784; GB:M36963; NID:G30072; PIDN:CAA31275.1; PID:G30073  
 R:Siolinen, R.; Huotari, M.; Hosikaka, S.L.; Prockop, D.J.; Trygvaason, K.  
 J. Biol. Chem. 263, 17217-17220, 1988  
 A:Title: The structural genes for alpha1 and alpha2 chains of human type IV collagen are  
 A:Reference number: A92690; MUID:89034231; PMID:3182844  
 A:Accession: B32117  
 A:Molecule type: DNA  
 A:Residues: 1-33 <SO11>  
 A:Cross-references: EMBL:J04217; EMBL:J05039; NID:G180759; PIDN:AAA53097.1; PID:G553233  
 R:Siolinen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Trygvaason, K.  
 J. Biol. Chem. 264, 13565-13571, 1989  
 A:Title: Structural organization of the gene for the alpha-1 chain of human type IV coll  
 A:Reference number: S16876; MUID:89340433; PMID:2701944  
 A:Accession: S16877  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-33 <SO12>  
 A:Cross-references: EMBL:J04217; NID:G180759; PIDN:AAA53097.1; PID:G553233  
 A>Note: this sequence was submitted to the EMBL Data Library, October 1988  
 R:Siolinen, R.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutzmann, R.; Kuehn, K.  
 Eur. J. Biochem. 168, 569-575, 1987  
 A:Title: Construction of a model for the aggregation and cross-linking region (7S domain  
 A:Reference number: S00165; MUID:88029476; PMID:3117548  
 A:Accession: S00165  
 A:Molecule type: protein  
 A:Residues: 37-247 <SIE1>  
 A>Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-Gly  
 R:Edle, U.A.; Goldik, R.; Mann, K.; Kuehn, K.  
 EMBO J. 12, 4795-4802, 1993  
 A:Title: The alpha-1-beta-1 integrin recognition site of the basement membrane collagen  
 A:Reference number: S39614; MUID:94038963; PMID:8223488  
 A:Accession: S39615  
 A:Molecule type: protein  
 A:Residues: 407-570 <EBL>  
 R:MacWright, R.S.; Benson, V.A.; Lovell, K.T.; van der Rest, M.; Fietzek, P.P.

Biochemistry 22, 4940-4948, 1983  
 A:Title: Isolation and characterization of pepsin-solubilized human basement membrane (C  
 A:Reference number: S16910; MUID:84053346; PMID:6416291  
 A:Accession: S16912  
 A:Molecule type: protein  
 A:Residues: 490-492, 'X', 494-496, 675-677, 'G', 679-680, 'G', 682, 684-685, 'P' <MAC>  
 R:Glanville, R.W.; Rauter, A.  
 Hoppe-Seyler's Z. Physiol. Chem. 362, 943-951, 1981  
 A:Title: Pepsin fragments of human placental basement-membrane collagens showing intern  
 A:Reference number: S16908; MUID:82005835; PMID:6792033  
 A:Accession: B58517  
 A:Molecule type: protein  
 A:Residues: 490-492, 'X', 494-501, 'P', 503-507, 952-957, 'X', 959-966, 'X', 968, 984-986, 'X', 988-  
 81-1195 <GLA>  
 R:Killem, P.D.; Francomano, C.A.; Yamada, Y.; Modi, W.S.; O'Brien, S.J.  
 Hum. Genet. 77, 318-324, 1987  
 A:Title: Partial structure of the human alpha-2(IV) collagen chain and chromosomal loca  
 A:Reference number: S01450; MUID:88085168; PMID:3692475  
 A:Accession: S01450  
 A:Molecule type: mRNA  
 A:Residues: 1040, 'L', 1042-1398, 'V', 1400-1418, 'W', 1420-1635, 'V', 1637-1712 <KIL>  
 A:Cross-references: EMBL:M24766; NID:G537328; PIDN:AAA52043.1; PID:G537329  
 R:Siebold, B.; Deutzmann, R.; Kuehn, K.  
 Eur. J. Biochem. 176, 617-624, 1988  
 A:Title: The arrangement of intra- and intermolecular disulfide bonds in the carboxyter  
 A:Reference number: S02550; MUID:89005112; PMID:2844531  
 A:Accession: S02550  
 A:Molecule type: protein  
 A:Residues: 1480-1535, 1545-1614, 1617-1662, 'H', 1664-1700, 'G', 1705-1708, 1710-1712 <SIE2>  
 A>Note: the sequence form Fig. 7 is inconsistent with that shown in Fig. 11 in having 1  
 R:Myers, U.C.; Howard, P.S.; Jelen, A.M.; Dion, A.S.; Macarak, E.J.  
 J. Biol. Chem. 262, 9231-9238, 1987  
 A:Title: Duplication of type IV collagen COOH-terminal repeats and species-specific exp  
 A:Reference number: A27114; MUID:87250571; PMID:2439508  
 A:Accession: B27114  
 A:Molecule type: mRNA  
 A:Residues: 1486-1574, 'I', 1576-1712 <MYE>  
 A:Cross-references: EMBL:J02760; NID:G180425; PIDN:AAA58422.1; PID:G180426  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C:Genetics:  
 A:Gene: GDB:COL4A2  
 A:Cross-references: GDB:119792; OMIM:120090  
 A:Map position: 13q34-13q34  
 A:Introns: 15/2; 33/3; 134/7; 1380/1; 1429/1; 1468/1; 1532/1; 1527/3 #status incomple  
 A>Note: the alpha 1(IV) and alpha 2(IV) chain genes are encoded on opposite strands with  
 C:Complex: Type IV collagen is a heterotrimer of two alpha 1(IV) chains (see PIR:CGU4B  
 domains (with disulfide and desmosine cross-links), dimeric associations among trimer c  
 rrupted helical domain (with disulfide and desmosine cross-links).  
 C:Function:  
 A:Description: structural component of basement membrane  
 C:Superfamily: collagen alpha 1(IV) chain  
 C:Keywords: basement membrane; cell binding; coiled coil; extracellular matrix; glycop  
 P1-28/Domain: signal sequence #status predicted <Sig>  
 F1-29-112/Product: collagen alpha 2(IV) chain predicted <Mat>  
 F1-29-57/Domain: amino-terminal nonhelical, NH1 <NH1>  
 F1-58-1485/Region: interrupted helical  
 F1-562-364/Region: cell attachment (R-G-D) motif  
 F1-784-786/Region: cell attachment (R-G-D) motif  
 F1-668-870/Region: cell attachment (R-G-D) motif  
 F1-689-891/Region: cell attachment (R-G-D) motif  
 F1-970-972/Region: cell attachment (R-G-D) motif  
 F1-1069-1071/Region: cell attachment (R-G-D) motif  
 F1-1228-1230/Region: cell attachment (R-G-D) motif  
 F1-1452-1454/Region: cell attachment (R-G-D) motif  
 F1-1486-1712/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
 F1-1495-1553/Domain: collagen IV carboxyl-terminal repeat <CT1>  
 F1-1603-1708/Domain: collagen IV carboxyl-terminal repeat <CT2>  
 F1-42, 47, 51, 53, 137, 483, 485/Disulfide bonds: interchain #status predicted  
 F1-47, 87, 90, 102, 165, 168, 225, 233, 242/Binding site: carboxylate (Lys) (covalent) #status  
 F1-57/Modified site: 5-hydroxylysine (Lys) #status atypical  
 F1-63, 75, 96, 114, 120, 123, 132, 150, 159, 186, 189, 198, 201, 213, 216, 219, 496, 499, 955, 964, 1103, 111

F:87-90,102,165,168,225,239,242/Modified site: 5-hydroxylysine (lys) #status experimental  
F:136/Binding site: carboxylate (asn) (covalent) #status experimental  
F:136/Modified site: 4-hydroxyproline (pro) #status atypical  
F:161-681/Disulfide bonds: #status predicted  
F:1275/Binding site: carboxylate (asn) (covalent) #status predicted  
F:1549-1590,1537-1559/Disulfide bonds: (or 1504-1593, 1537-1590) #status experimental  
F:1549-1555,1658-1665/Disulfide bonds: #status experimental  
F:1612-1705,1646-1708/Disulfide bonds: (or 1612-1708, 1646-1705) #status experimental

Query Match 41.5%; Score 44; DB 1; Length 1712;  
Best Local Similarity 63.6%; Pred. No. 2.1e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AIEGPTLRQWL 12  
: || || || ||  
Db 7 AVAGPALRRL 17

RESULT 24

GNMVVY  
genome polypeptide - West Nile virus  
N:Contains: core protein V2; membrane-associated glycoprotein NV2 precursor; membrane-as  
sor; nonstructural protein NV5  
C:Species: West Nile virus  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004  
C:Accession: A25256  
R:Caetle, E.; Leiden, U.; Nowak, T.; Wengler, G.; Wengler, G.  
Virology 149, 10-26, 1986  
A:Title: Primary structure of the West Nile flavivirus genome region coding for all non  
A:Reference number: A25256; MUID:86124703; PMID:3753811  
A:Accession: A25256  
A:Molecule type: genomic RNA  
A:Reidsites: 1-3430 <CNS>  
A:Cross-references: UNIPROT:P06935; GB:M10103; GB:M12294; NID:g336167; PIDN:AAA48498.1;  
A:Note: Parts of this sequence, including the amino ends of the mature proteins, were de  
C:Superfamily: Yellow fever virus genome polypeptide  
C:Keywords: ATP; core protein; glycoprotein; membrane-associated protein; nucleotide bind  
F:1-32/Product: core protein V2 #status predicted <V2>  
F:105-233/Product: membrane-associated glycoprotein NV2 precursor #status predicted <NV2>  
F:105-123/Domains: nonterminal signal sequence #status predicted <2S>  
F:124-233/Product: membrane-associated glycoprotein NV2 #status predicted <2NV>  
F:216-233/Product: membrane-associated nonglycosylated protein V1 #status predicted <NV1>  
F:275-787/Product: membrane-associated glycoprotein V3 precursor #status predicted <NV3>  
F:275-290/Domains: nonterminal signal sequence #status predicted <3S>  
F:291-787/Product: membrane-associated glycoprotein V3 #status predicted <3NV>  
F:788-2109/Product: nonstructural protein NV4 #status predicted <NV4>  
F:1695-1702/Region: nucleotide-binding motif A (P-loop)  
F:1782-1787/Region: nucleotide-binding motif B  
F:1786-1789/Region: DEAH motif  
F:1780-3427/Product: nonstructural protein NV5 #status predicted <NV5>  
F:138-917,962,994,1289,1659,1666,2336,2459,2489,2573,2739,2759,2864,2902/Binding site: C

Query Match 41.5%; Score 44; DB 1; Length 3430;  
Best Local Similarity 46.7%; Pred. No. 4.5e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EGPTRLQWLHNGRD 18  
: || || || || ||  
Db 3169 KGPXRTWLFWENGEE 3183

RESULT 25

GNMVVY  
genome polypeptide - Kunjin virus (strain MRM61C)  
N:Contains: capsid protein C; envelope protein E; membrane protein M; nonstructural protein  
in NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: Kunjin virus  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C:Accession: A28697  
R:Coia, G.; Parker, M.D.; Speight, G.; Byrne, M.E.; Westaway, E.G.  
J. Gen. Virol. 69, 1-21, 1988  
A:Title: Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene c  
A:Reference number: A28697; MUID:88089524; PMID:2826655

A:Accession: A28697  
A:Molecule type: genomic RNA  
A:Residues: 1-3433 <COT>  
A:Cross-references: UNIPROT:P14335; GB:D00246; NID:g221966; PION:BAA00176.1; PID:g221967  
C:Superfamily: yellow fever virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; membrane protein; nonstructural prote  
F:1-123/Product: capsid protein C #status predicted <CP>  
F:124-280/Product: membrane protein M precursor #status predicted <MP>  
F:124-215/Domain: nonterminal signal sequence #status predicted <SIG>  
F:126-280/Product: membrane protein M #status predicted <MP>  
F:1291-791/Product: envelope protein E #status predicted <EP>  
F:792-1143/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1144-1374/Product: nonstructural protein NS2a #status predicted <N2a>  
F:1135-1505/Product: nonstructural protein NS2b #status predicted <N2b>  
F:1506-2124/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1699-1706/Region: nucleotide-binding motif A (P-loop)  
F:1786-1791/Region: nucleotide-binding motif B  
F:1790-1793/Region: DEAH motif  
F:2125-2273/Product: nonstructural protein NS4a #status predicted <NA4>  
F:2274-2528/Product: nonstructural protein NS4b #status predicted <N4b>  
F:2529-3433/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 41.5%; Score 44; DB 1; Length 3433;  
Best Local Similarity 46.7%; Pred. No. 4.5e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 3172 KGPRTVLTSENGE 3186

RESULT 26

AG0147  
Probable membrane protein YPO1203 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Dates: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AG0147  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tildall, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AG0147  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-296 <KOR>  
A:Cross-references: UNIPROT:Q8ZGS7; GB:AL590842; PION:CAC90042.1; PID:G15979263; GSPDB:C  
C:Genetics:  
A:Gene: YPO1203

Query Match 41.0%; Score 43.5; DB 2; Length 296;  
Best Local Similarity 71.4%; Pred. No. 37;  
Matches 10; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 LAIRG---PTLRW 11  
||| |||||  
Db 58 LAIRGHAPTLRW 71

RESULT 27

H82539  
Protein-L-isopartate O-methyltransferase XF2585 [imported] - Xylella fastidiosa (stra  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 12-Jul-2004  
C:Accession: H82539  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82539  
A>Status: preliminary



A:Molecule type: DNA  
A:Residues: 1-218 <SIM>  
A:Cross-references: UNIPROT:Q9PAD3; GB:AE004065; GB:AE003849; NID:G9107795; PIND:AA6538  
A:Experimental source: strain 9abc  
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
B:Ribeiro, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carreir, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Klitajma, J.P.; Krieger, J.E.; Kurrae, E.E.; Laig  
chido, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XR2585  
C:Superfamily: Escherichia coli protein-L-isoaspartate(D-aspartate) O-methyltransferase  
Query Match 40.6%; Score 43; DB 2; Length 218;  
Best Local Similarity 87.5%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 10 QWLGNGR 17  
Db 164 QWLGNGR 171  
RESULT 28  
A:Accession: A45822  
beta-lactamase (EC 3.5.2.6) precursor - Streptomyces badius  
C:Species: Streptomyces badius  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-2004  
C:Accession: A45822  
R:Forstman, M.; Haegsaetrom, B.; Lindgren, L.; Jaurin, B.  
J. Gen. Microbiol. 136, 589-598, 1990  
A:Title: Molecular analysis of beta-lactamases from four species of Streptomyces: compar  
A:Reference number: A45822; MUID:90362045; PMID:2391494  
A:Accession: A45822  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <FOR>  
A:Cross-references: UNIPROT:P35391; GB:M34178; NID:G153182; PIND:AAA26707.1; PID:G153183  
C:Superfamily: Beta-lactamase I  
C:Keywords: hydrolase  
F:93/Active site: Ser #status predicted  
Query Match 40.6%; Score 43; DB 2; Length 313;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
QY 4 EGPTLROMLHNGRDT 19  
Db 187 EBPGLSRWPGKXKDT 202  
RESULT 29  
B69901  
fatty-acid desaturase homolog yocE - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
C:Accession: B69901  
R:Kunze, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cht  
A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier  
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullio, M.F.  
Koetter, P.; Koningsreid, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
A:Authors: Schlaich, S.; Schoeber, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero  
kouchi, M.; Tanakoshi, A.; Tanaka, T.; Terpeira, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumelein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: B69901  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-352 <KUN>  
A:Cross-references: UNIPROT:Q34653; GB:Z99114; GB:AL009126; NID:G2634230; PIND:CA813810  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yocE  
C:Superfamily: Fatty acid (acyl-CoA) desaturase  
Query Match 40.6%; Score 43; DB 2; Length 352;  
Best Local Similarity 50.0%; Pred. No. 53;  
Matches 10; Conservative 1; Mismatches 3; Indels 6; Gaps 1;  
QY 2 AIEG-----PTLRQMLHGN 15  
Db 251 AVEGSSFYRLPRLQLWLTGN 270  
RESULT 30  
T05270  
probable serine/threonine-specific protein kinase (EC 2.7.1.-) T4L20.80 - Arabidopsis ti  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05270  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysheart, C.; Desseville, R.; De Clerck, R.; De  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, September 1998  
A:Reference number: Z15406  
A:Accession: T05270  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-481 <BEV>  
A:Cross-references: UNIPROT:Q65676; EMBL:AL023094  
A:Experimental source: cultivar Columbia; BAC clone T4L20  
C:Genetics:  
A:Map position: 4  
A:Intons: 179/2; 216/2; 257/2; 314/2; 357/3  
A:Note: T4L20.80  
C:Keywords: phosphotransferase  
Query Match 40.6%; Score 43; DB 2; Length 481;  
Best Local Similarity 53.8%; Pred. No. 75;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 3 IEQPTLROMLHGN 15  
Db 223 IDWGNLROMLHGN 235  
RESULT 31  
G89894  
protein kinase [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: G89894  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.  
Lancet 357, 1225-1240, 2001  
C.; Shib, T.; Hatori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: G89894  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <KOR>

A:Cross-references: UNIPROT:Q99UP8; GB:BA000018; PID:g13701020; PIDN:BA842315.1; GSPDB:6  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA1063

Query Match 40.6%; Score 43; DB 2; Length 664;  
Best Local Similarity 50.0%; Pred. No. 1,1e+02;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 IEPTLRWLHGN 16  
|||  
Db 90 IEPTLRWLHGN 103

## RESULT 32

A43254  
protein-tyrosine-phosphatase (EC 3.1.3.48) corkscrew - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A43254  
R:Perkins, L.A.; Larsen, I.; Perrimon, N.  
Cell 70, 225-236, 1992  
A>Title: corkscrew encodes a putative protein tyrosine phosphatase that functions to tra  
A:Reference number: A43254; PMID:92346711; PMID:1638629  
A:Accession: A43254  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-841 <PER>  
A:Cross-references: UNIPROT:P29349; GB:M94730; NID:g157144; PID:g157145  
A:Experimental source: embryo  
A>Note: sequence extracted from NCBI backbone (NCBIN:109964, NCBI:109965)  
C:Genetics:  
A:Gene: FlyBase:cw  
A:Cross-references: FlyBase:FBgn0000382  
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosph  
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase  
F:6-101/Domain: SH2 homology <SH2>  
F:111-203/Domain: SH2 homology <SH2B>  
F:552-634/Domain: protein-tyrosine-phosphatase homology <PTP>  
F:583/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:589/Binding site: substrate phosphate (Arg) #status predicted

Query Match 40.6%; Score 43; DB 2; Length 841;  
Best Local Similarity 60.0%; Pred. No. 1,4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 PTLRWLHGN 15  
|||  
Db 106 PTLRWLHGN 115

## RESULT 33

E71376  
conserved hypothetical protein TP0025 - syphilis spirochete  
C:Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: E71376  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, G.G.; Dodson, R.; Gwin  
rson, J.; Khakh, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utecherback, T.; McD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 261, 375-388, 1998  
A>Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A:Reference number: A71250; PMID:98332770; PMID:9665876  
A:Accession: E71376  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1023 <COU>  
A:Cross-references: UNIPROT:Q83059; GB:AE001187; GB:AE000520; NID:g3322273; PIDN:AA6501  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0025

Query Match 40.6%; Score 43; DB 2; Length 1023;

Best Local Similarity 37.5%; Pred. No. 1,7e+02;  
Matches 9; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

QY 4 EGP-----TLRWLHGNRDT 19  
|||  
Db 396 DGPSFLVMORSLRGLHGNAPES 419

## RESULT 34

E82797  
conserved hypothetical protein XF0501 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C:Species: *Xylella fastidiosa*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: E82797  
R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: A82515; PMID:20365177; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: E82797  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-104 <SIM>  
A:Cross-references: UNIPROT:Q9PG01; GB:AE003899; GB:AE003849; NID:g9105351; PIDN:AAF8331  
A:Experimental source: strain 9a5c  
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briotes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawaak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tuhache, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0501

Query Match 39.6%; Score 42; DB 2; Length 104;  
Best Local Similarity 42.9%; Pred. No. 20;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAIEPTLRWLHGN 14  
|||  
Db 30 LGVSPFLVMORSLRGLHGNAPES 43

## RESULT 35

AC3086  
sarcosine oxidase delta subunit [imported] - *Agrobacterium tumefaciens* (strain C58, Dupo  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AC3086  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, L  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, W.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; PMID:21608550; PMID:11743193  
A:Accession: AC3086  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-106 <KUN>  
A:Cross-references: UNIPROT:Q8U7Y8; GB:AE000869; PIDN:AA45105.1; PID:g17742774; GSPDB:6  
A:Experimental source: strain C58 (Dugont)  
C:Genetics:  
A:Gene: soxH



A:Map position: linear chromosome  
C:Superfamily: Corynebacterium sp. sarcosine oxidase delta chain

Query Match 39.6%; Score 42; DB 2; Length 106;  
Best Local Similarity 47.1%; Pred. No. 20;

Matches 8; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 3 IEPTLRQW-LHGNGR 17  
DB 50 VKGPHFRWRHLHGCR 66

RESULT 36  
P98200  
sarcosine oxidase delta chain (sarcosine oxidase chain d) [imported] - Agrobacterium tum

C:Species: Agrobacterium tumefaciens  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C:Accession: F98200  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: F98200  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-106 <RUR>

A:Cross-references: UNIPROT:Q8U7Y8; GB:AE007870; PIDN:AAK89128.1; PID:G15158936; GSPDB:G

C:Genetics:  
A:Gene: AGR\_1105  
A:Map position: linear chromosome

C:Superfamily: Corynebacterium sp. sarcosine oxidase delta chain

Query Match 39.6%; Score 42; DB 2; Length 106;  
Best Local Similarity 47.1%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 3 IEPTLRQW-LHGNGR 17  
DB 50 VKGPHFRWRHLHGCR 66

RESULT 37  
D83161  
hypothetical protein PA3885 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C:Accession: D83161  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Iarbig, K.; Lim,  
J.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; PMID:20437337; PMID:10984043

A:Accession: D83161  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-218 <STO>  
A:Cross-references: UNIPROT:Q8HXC7; GB:AE004805; GB:AE004091; NID:9950055; PIDN:AA0727

A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3885

Query Match 39.6%; Score 42; DB 2; Length 218;  
Best Local Similarity 57.1%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGPTLRQWLHGNGR 17  
DB 125 EGPVLMHCKGNR 138

RESULT 38

T44657  
protein GP80 [imported] - bovine herpesvirus 4

C:Species: Bovine herpesvirus 4  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T44657  
R:Monte, P.; Van Santen, V.L.; Filie, P.; Lyaku, J.R.; Bublott, M.; Pastoret, P.; Thiry  
submitted to the EMBL Data Library, February 1997  
A:Description: Identification and characterization of bovine herpesvirus 4 GP80: a new g

A:Reference number: Z22823  
A:Accession: T44657  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-273 <LOW>

A:Cross-references: UNIPROT:P87519; EMBL:Z84818; PIDN:CA06616.1  
C:Genetics:  
A:introns: 177/1

Query Match 39.6%; Score 42; DB 2; Length 273;  
Best Local Similarity 61.5%; Pred. No. 58;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 PTLRQWLHGNGRD 18  
DB 210 PTRKRVILHGNGFD 222

RESULT 39  
B72053  
glyceralddehyde 3-phosphate dehydrogenase CP0123 [imported] - Chlamydia pneumoniae (

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: B72053; G81613  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.  
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; PMID:99206606; PMID:10192388

A:Accession: B72053  
A:Molecule type: DNA  
A:Residues: 1-335 <ARN>

A:Cross-references: UNIPROT:Q9Z7T0; GB:AE001647; GB:AE001363; NID:94376920; PIDN:AA0187  
A:Experimental source: strain CWD029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; PMID:20150255; PMID:10684935

A:Accession: G81613  
A:Molecule type: DNA  
A:Residues: 1-335 <REA>

A:Cross-references: GB:AE002173; GB:AE002161; NID:97189033; PIDN:AAF38006.1; PID:971890  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: gAPA; CP0123  
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 39.6%; Score 42; DB 2; Length 335;  
Best Local Similarity 37.5%; Pred. No. 72;  
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LATEPTLRQWLHGNG 16  
DB 185 LVVDGSPSKDWRGGRG 200

RESULT 40  
E86568  
glyceralddehyde-3-P dehydrogenase [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86568  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

```

A:Reference number: AB6491; MUID:20303049; PMID:10871362
A:Accession: BG6568
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <STO>
A:Cross-references: UNIPROT:Q9Z7T0; GB:BA000008; NID:g8978996; PIDN:BA98831.1; GSPDB:GN
A:Experimental source: strain UJ38
A:Genetics:
A:Gene: gapA
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match          39.6%; Score 42; DB 2; Length 335;
Best Local Similarity 37.5%; Pred. No. 72;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY      1 LAIEGTLRQWLHNGG 16
        ::|||::|||
Db       185 LVVDGSKDWRGGRG 200

RESULT 41
S43339
Glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - red alga (Chlorella
C:Species: Chondrus crispus (carrageen)
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S43339; S32692
C:Residues: 1-335 <LIA>
C:Cross-references: UNIPROT:P34920; EMBL:X73036; NID:G440394; PIDN:CA51517.1; PID:G440394
R:Liud, M.F.; Valentín, C.; Brandt, U.; Bouget, F.Y.; Kloareg, B.; Cerff, R.
Submitted to the EMBL Data Library, April 1993
A:Description: The evolutionary origin of red algae as deduced from the nuclear genes encoding
A:Reference number: S43339; MUID:94083567; PMID:8260635
A:Accession: S43339
A:Molecule type: DNA
A:Residues: 1-335 <LIA>
A:Cross-references: UNIPROT:P34920; EMBL:X73036; NID:G440394; PIDN:CA51517.1; PID:G440394
R:Liud, M.F.; Valentín, C.; Martin, W.; Bouget, F.Y.; Kloareg, B.; Cerff, R.
Submitted to the EMBL Data Library, April 1993
A:Description: The evolutionary origin of red algae as deduced from the nuclear genes encoding
A:Reference number: S32692
A:Accession: S32692
A:Molecule type: mRNA
A:Residues: 1-249, 'A', 251-335 <LIM>
A:Cross-references: EMBL:X73034; NID:G297453; PIDN:CA51515.1; PID:G297454
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match          39.6%; Score 42; DB 2; Length 335;
Best Local Similarity 40.0%; Pred. No. 72;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY      3 IEGLTLRQWLHNGR 17
        ::|||::|||
Db       188 VDGSKDWRGGRG 202

RESULT 42
A30754
Hypochemical protein tera - Alcaligenes sp.
C:Species: Alcaligenes sp.
C:Date: 19-May-1989 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: J0361; A30754
R:Jobling, M.G.; Ritchie, D.A.
Gene 66, 245-258, 1988
A:Title: Nucleotide sequence of a plasmid determinant for resistance to cellulium anionomans
A:Reference number: J0361; MUID:89006266; PMID:3049247
A:Accession: J0361
A:Molecule type: DNA
A:Residues: 1-339 <UOI>
A:Cross-references: UNIPROT:Q44313
R:Jobling, M.G.
submitted to GenBank, September 1988
A:Reference number: A30754
A:Accession: A30754
A:Molecule type: DNA

```

A;Residues: 1-224,'M',226-339 <J02>

```
Query Match      39.6%; Score 42; DB 2; Length 339;  
Best Local Similarity 47.1%; Pred. No. 73;  
Matches 8; Conservative 4; Mismatches 1; Indels 4; Gaps 1;  
  
OY          2 AIEGPTLRQWLHGNGRD 18  
              |||::|||  
Db          234 ALDG-----EWHLHNGRE 246
```

RESULT 43

A86298

hypothetical protein F309\_10 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: A86298

R;Thellogis: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huitzar, L.  
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L., Jenkins, J., Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, U.S.; Matti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; PMID:21016719; PMID:11130712

A;Accession: A86298

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-407 <STO>

A;Cross-references: UNIPROT:Q9SAZ9; GB:AEO05172; NID:g966351; PIDN:AAD34682.1; GSPDB:GN  
C;Genetics:

A;Map position: 1

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

```
Query Match      39.6%; Score 42; DB 2; Length 407;  
Best Local Similarity 35.7%; Pred. No. 89;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
  
OY          3 IEIPTLRQWLHGNG 16  
              ::|||::|  
Db          257 VDGPMDKMRGRG 270
```

RESULT 44

P96826

hypothetical protein T8K14.5 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: P96826

R;Theologis: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huitzar, L.  
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, U.S.; Matti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; PMID:21016719; PMID:11130712

A;Accession: P96826

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-422 <STO>

A;Cross-references: UNIPROT:Q9SAJ6; GB:AEO05173; NID:g94835756; PIDN:AAD30223.1; GSPDB:GN  
C;Genetics:

A;Gene: T8K14.5

A;Map position: 1

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 39.6%; Score 42; DB 2; Length 422;  
 Best Local Similarity 35.7%; Pred. No. 93;  
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 IEQPTLRQWLHGNG 16  
 ::|||::|||  
 DB 272 VDGPSMKDWRGGRG 285

## RESULT 45

S51837  
 glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) precursor - Sc  
 N:Alternate names: PPSD17 protein  
 C:Species: pinus blyvestris (Scotch pine)  
 C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S51837  
 R:Meyer-Gauen, G.; Schnarrenberger, C.; Cerff, R.; Martin, W.  
 Plant Mol. Biol. 26, 1155-1166, 1994  
 A>Title: Molecular characterization of a novel, nuclear-encoded, NAD(+)-dependent glycer  
 A:Reference number: S51836; MUID:95111098; PMID:7811973  
 A:Accession: S51837  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <MBY>  
 A:Cross-references: UNIPROT:Q37264; EMBL:L32561; NID:G1100224; PTDN:AAD10214.1; PID:G110  
 C:Genetics:  
 A:Genome: nuclear  
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase  
 C:Keywords: chloroplast; gluconeogenesis; glycolysis; oxidoreductase

Query Match 39.6%; Score 42; DB 2; Length 433;  
 Best Local Similarity 35.7%; Pred. No. 95;  
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 IEQPTLRQWLHGNG 16  
 ::|||::|||  
 DB 283 VDGPSMKDWRGGRG 296

Search completed: September 1, 2005, 16:22:59  
 Job time : 17.4892 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 70.6691 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-11  
Perfect score: 106  
Sequence: 1 LAIBEPYLRQWHLHGNGRDT 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	52.4	497	2	06LKY2 photobacter
2	51	48.1	71	2	076H55
3	51	48.1	580	2	089RH2
4	51	48.1	667	2	08CSV9
5	51	48.1	690	2	082RY5
6	50	47.2	326	2	06XK8
7	50	47.2	407	2	08XYS6
8	50	47.2	486	2	0877P6
9	50	47.2	486	2	0877Q7
10	50	47.2	486	2	0877V4
11	50	47.2	690	1	EP42_HUMAN
12	50	47.2	690	1	EP42_MOUSE
13	50	47.2	691	2	06NXX8
14	50	47.2	973	2	08YPC5
15	50	47.2	1774	1	MSAS_PENPA
16	49	46.2	245	2	066272
17	49	46.2	249	2	082989
18	49	46.2	278	2	09XDY0
19	49	46.2	412	2	088D08
20	49	46.2	419	1	EREB_ECOLI
21	49	46.2	547	2	07SFG9
22	48.5	45.8	88	2	06YSI5
23	48	45.3	154	2	082S29
24	48	45.3	168	2	09V492
25	48	45.3	392	2	0688F0
26	48	45.3	396	2	067U70
27	48	45.3	631	2	06LP74
28	48	45.3	686	1	EP42_BOVIN
29	48	45.3	687	2	046509
30	48	45.3	1916	2	084PP9
31	47	44.3	182	2	075174

32	47	44.3	344	2	055838
33	47	44.3	349	1	MOA1_CAUCR
34	47	44.3	391	2	09HX67
35	47	44.3	483	2	0886U0
36	47	44.3	483	2	076676
37	47	44.3	3425	2	07F918
38	46.5	43.9	495	2	07MP20
39	46.5	43.9	600	2	093IU2
40	46.5	43.9	602	2	P72407
41	46	43.4	85	2	08RSL6
42	46	43.4	256	2	09NBS3
43	46	43.4	279	1	06H5Y4
44	46	43.4	349	1	MOA1_RHME
45	46	43.4	434	2	074061
46	46	43.4	451	2	087IX8
47	46	43.4	495	2	08E9K7
48	46	43.4	1044	2	08BDI0
49	46	43.4	1154	2	08BNE9
50	46	43.4	4190	2	083Y48
51	45.5	42.9	333	1	CBRR_XANFL
52	45	42.5	81	2	09NDL7
53	45	42.5	139	2	09G5F9
54	45	42.5	194	2	086479
55	45	42.5	230	2	088H66
56	45	42.5	235	2	09KD40
57	45	42.5	245	2	082987
58	45	42.5	302	1	PHER_PSEUE
59	45	42.5	304	2	097V63
60	45	42.5	346	2	08P199
61	45	42.5	346	2	09A045
62	45	42.5	346	2	07CNA2
63	45	42.5	348	1	RMUB_STRMU
64	45	42.5	348	2	08GTF9
65	45	42.5	348	2	08DZB1
66	45	42.5	348	2	08E4X2
67	45	42.5	349	2	07W0B9
68	45	42.5	349	2	07WCC0
69	45	42.5	349	2	07WOC4
70	45	42.5	379	2	08PRC6
71	45	42.5	401	2	096C72
72	45	42.5	403	2	088NU2
73	45	42.5	429	2	07UTX3
74	45	42.5	467	2	08H846
75	45	42.5	492	2	022764
76	45	42.5	577	2	08XYA0
77	45	42.5	639	1	P2B1_CRYNV
78	45	42.5	641	2	09Y879
79	45	42.5	654	2	09AVL8
80	45	42.5	656	2	06YZ49
81	45	42.5	794	2	09U353
82	45	42.5	821	2	0924E2
83	45	42.5	1082	2	063UG8
84	45	42.5	1382	1	IF3A_HUMAN
85	45	42.5	2042	2	09VUG3
86	45	42.5	4163	2	09LAE6
87	44.5	42.0	202	2	06AB11
88	44.5	42.0	244	2	07W0F8
89	44.5	42.0	244	2	07WP46
90	44.5	42.0	319	2	09RKM5
91	44.5	42.0	513	2	088759
92	44.5	42.0	513	2	09B857
93	44.5	42.0	513	2	09B857
94	44	41.5	81	2	09NDL5
95	44	41.5	81	2	09NDL5
96	44	41.5	81	2	09NDL9
97	44	41.5	89	2	09WUG8
98	44	41.5	109	2	06AP30
99	44	41.5	176	2	0656A5
100	44	41.5	192	2	06U027

## ALIGNMENTS

055838	yokose viru	09ac48	caulobacter
09ac48	caulobacter	09hx67	pseudomonas
09hx67	pseudomonas	0886U0	oryza sativ
0886U0	oryza sativ	076676	oryza sativ
076676	oryza sativ	07F918	yokose viru
07F918	yokose viru	07MP20	vibrio vuln
07MP20	vibrio vuln	093IU2	streplococce
093IU2	streplococce	P72407	streplococce
P72407	streplococce	08RSL6	uncultured
08RSL6	uncultured	09NBS3	tachypleus
09NBS3	tachypleus	06H5Y4	oryza sativ
06H5Y4	oryza sativ	092PB4	rhizobium m
092PB4	rhizobium m	074061	cenarchaeum
074061	cenarchaeum	087IX8	vibrio para
087IX8	vibrio para	08E9K7	shewanella
08E9K7	shewanella	08BDI0	synecococce
08BDI0	synecococce	08BNE9	homo sapien
08BNE9	homo sapien	083Y48	pseudomonas
083Y48	pseudomonas	P25545	xanthobacte
P25545	xanthobacte	09nd17	hydra magni
09nd17	hydra magni	09g5f9	hydra atten
09g5f9	hydra atten	086479	streplococce
086479	streplococce	088H66	pseudomonas
088H66	pseudomonas	09KD40	bacillus ha
09KD40	bacillus ha	082987	erythroba
082987	erythroba	P31019	pseudomonas
P31019	pseudomonas	097V63	sulfolobus
097V63	sulfolobus	08P199	streplococce
08P199	streplococce	09A045	streplococce
09A045	streplococce	07CNA2	streplococce
07CNA2	streplococce	P95780	streplococce
P95780	streplococce	08g1p9	streplococce
08g1p9	streplococce	08db11	streplococce
08db11	streplococce	08e4x2	streplococce
08e4x2	streplococce	07w0b9	bordetella
07w0b9	bordetella	07wcc0	bordetella
07wcc0	bordetella	07woc4	bordetella
07woc4	bordetella	08prc6	xanthomonas
08prc6	xanthomonas	096c72	homo sapien
096c72	homo sapien	088nu2	pseudomonas
088nu2	pseudomonas	07uxx3	rhodospirell
07uxx3	rhodospirell	08b46	oryza sativ
08b46	oryza sativ	022764	arabidopsis
022764	arabidopsis	08xya0	raibetonia s
08xya0	raibetonia s	042773	cryptococcu
042773	cryptococcu	09y879	cryptococcu
09y879	cryptococcu	09av18	oryza sativ
09av18	oryza sativ	06yz49	oryza sativ
06yz49	oryza sativ	09u353	caenorhabd
09u353	caenorhabd	0924e2	apicomplexa
0924e2	apicomplexa	063ug8	burkholderi
063ug8	burkholderi	01452	homo sapien
01452	homo sapien	09v193	drosophila
09v193	drosophila	09lae6	rhizobium 1
09lae6	rhizobium 1	06ab11	propionibac
06ab11	propionibac	07w0f8	bordetella
07w0f8	bordetella	07wp46	bordetella
07wp46	bordetella	09rkms	streplococce
09rkms	streplococce	099044	tarsius ban
099044	tarsius ban	088759	tarsius ban
088759	tarsius ban	09b857	tarsius ban
09b857	tarsius ban	09nd15	tina formos
09nd15	tina formos	09nd19	hydractinia
09nd19	hydractinia	09wug8	mus musculu
09wug8	mus musculu	06ap30	leifsonia x
06ap30	leifsonia x	0656a5	oryza sativ
0656a5	oryza sativ	06u027	oryza sativ

```

RESULT 1
ID 06LKY2; PRELIMINARY; PRT; 497 AA.
AC 06LKY2;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein YP00843.
GN Name=YP00843; OrderedLocustNames=PBPRB0156;
OS Photobacterium profundum (Photobacterium SP. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cesarro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378675; CAG22029.1; -.
KW Complete proteome.
SQ SEQUENCE 497 AA; 57026 MW; B977EF6C53465289 CRC64;

Query Match 52.4%; Score 55.5; DB 2; Length 497;
Best Local Similarity 52.2%; Pred. No. 3.5;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy 2 AIEG-----PTLRQWLHNGRDT 19
Db 104 AMEGSRVLPFLAAMHLANGQVT 126

RESULT 2
ID 076H55; PRELIMINARY; PRT; 71 AA.
AC 076H55;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Cto.
GN Name=cro;
OS Salmoneilla typhimurium bacteriophage ST104.
OC Viruses.
OX NCBI_TaxID=221029;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15071057;
RA Tanaka, K., Nishimori T., Kanno T.,
RA Ishihara R., Sameshima T., Akiba M., Nakazawa M., Yokomizo Y.,
RA Uchida I.;
RT "Molecular characterization of a prophage of Salmoneilla enterica
RT serotype Typhimurium DT104."
RL J. Clin. Microbiol. 42:1807-1812 (2004).
DR EMBL; AB102868; BAD15187.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fis
DR InterPro; IPR009061; Putativ_DNA_bind.
DR PRINTS; PR01590; HTHFIS.
SQ SEQUENCE 71 AA; 7484 MW; 577499E7BFP9EAD0 CRC64;

Query Match 48.1%; Score 51; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 IEGPTLRQWLHNGRDT 18
Db 25 VKOPTYRWMLHGGGID 40

```

```

RESULT 3
ID 089RH2; PRELIMINARY; PRT; 580 AA.
AC 089RH2;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE B112800 protein.
GN OrderedLocustNames=b112800;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiales; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005945; BAC48065.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000719; PP2C_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00481; PP2C; 1.
DR Prodom; PD000001; Proc_kinase; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 580 AA; 64916 MW; 6AD3A06E6FAE143B CRC64;

Query Match 48.1%; Score 51; DB 2; Length 580;
Best Local Similarity 46.7%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 IEGPTLRQWLHNGR 17
Db 355 IEGQTLROWMLDNPR 369

RESULT 4
ID 08CSV9; PRELIMINARY; PRT; 667 AA.
AC 08CSV9;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protein kinase.
GN OrderedLocustNames=SE0895;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228)."
Mol. Microbiol. 49:1577-1593 (2003).

```

CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL: AB016746; AAO04492.1; -.  
 DR HSSP: P71584; 1067.  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0008658; P:penicillin binding; IEA.  
 DR GO: GO:000674; P:protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR005543; PASTA.  
 DR InterPro: IPR000719; Prot Kinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam: PF03793; PASTA\_2.  
 DR Pfam: PF00069; PKinase; 1.  
 DR ProDom: PD000001; Prot\_Kinase; 1.  
 DR SMART: SM00740; PASTA; 3.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR KMW: ATP-binding; Complete proteome; Kinase;  
 KMW: Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 667 AA; 7541 MW; 47987784531CDD97 CRC64;

Query Match 48.1%; Score 51; DB 2; Length 667;  
 Best Local Similarity 57.1%; Pred. No. 25;  
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 IEGPTLRQWLHGNG 16  
 DB 90 IEGPTLAETIYHSHG 103

RESULT 5  
 O8XRYS PRELIMINARY; PRT; 690 AA.  
 AC O8XRYS;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative transposase.  
 GN OrderedLocustNames=SAV7;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NC NCB1\_Taxid=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osomoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 DR EMBL: AP005021; BAC67716.1; -.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0015074; P:DNA integration; IEA.  
 DR GO: GO:0006310; P:DNA recombination; IEA.  
 DR InterPro: IPR011010; DNA\_brx\_join\_enz.  
 DR InterPro: IPR002104; Phage\_integrase.

DR Pfam: PF00589; Phage\_integrase; 1.  
 KMW Complete proteome.  
 SQ SEQUENCE 690 AA; 77303 MW; 867C6DB9A390EA91 CRC64;

Query Match 48.1%; Score 51; DB 2; Length 690;  
 Best Local Similarity 52.4%; Pred. No. 26;  
 Matches 11; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 5 GPTLRQW-----LHGNGRDT 19  
 DB 486 GETLRDWDSTPHLHGEGTDT 506

RESULT 6  
 O69XG8 PRELIMINARY; PRT; 326 AA.  
 ID O69XG8;  
 AC O69XG8;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein OSUNBD0066M12.35.  
 GN Name=OSUNBD0066M12.35;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 NC NCB1\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC  
 RT clone:OSUNBD0066M12.";  
 RL Submitted (NOV-2002) to the EMBL/Genbank/DDJ databases.  
 DR EMBL: AP005916; BAC36572.1; -.  
 KMW Hypothetical protein.  
 SQ SEQUENCE 326 AA; 36299 MW; A4FAF7EFLAB6DD9B CRC64;

Query Match 47.2%; Score 50; DB 2; Length 326;  
 Best Local Similarity 80.0%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 RQWLHGNGRD 18  
 DB 33 RQWLHGNGDD 42

RESULT 7  
 O8XYS6 PRELIMINARY; PRT; 407 AA.  
 ID O8XYS6;  
 AC O8XYS6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE PROBABLE PHAGE PH1-105 ORF25-LIKE PROTEIN.  
 GN Name=RS04076; OrderedLocustNames=RSC1682;  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 NC NCB1\_Taxid=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SM111000;  
 RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlet M., Billault A., Brotier P., Camus J.C., Cactolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,  
 RA Caspin C., Davie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Signier P., Thebaud P., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646065; CAD15384.1; -.  
 DR Pfam: PF04860; Phage\_portal; 1.

DR TIGRfams; TIGR01537; portal\_HK97; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 407 AA; 45951 MW; CC2236A78E65A06C CRC64;

Query Match  
 Best Local Similarity 47.2%; Score 50; DB 2; Length 407;  
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 IEGPLTROMLHGNGRDT 19  
 :|:|||||:  
 Db 20 LTGRNLQEWLHGDCGAT 36

RESULT 8  
 ID 0877P6 PRELIMINARY; PRT; 486 AA.  
 AC 0877P6;  
 DT 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
 DE ISPy6, transposase.  
 GN OrderedlocusNames=PSPT00968, PSPT01418, PSPT01439, PSPT02204,  
 PSPT04792;  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Davidse S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidse S.C., Brinkac L.M., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
 RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinour S., Chatterjee A.K.,  
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,  
 RA Bender C.L., White O., Fraser C.M., Collier A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 Pseudomonas syringae pv. tomato DC3000."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).

DR EMBL; AE016859; AA054502.1; -;  
 DR EMBL; AE016860; AA054939.1; -;  
 DR EMBL; AE016860; AA054960.1; -;  
 DR EMBL; AE016863; AA055720.1; -;  
 DR EMBL; AE016873; AA058222.1; -;  
 DR TIGR; PSPT00968; -;  
 DR TIGR; PSPT01418; -;  
 DR TIGR; PSPT01439; -;  
 DR TIGR; PSPT02004; -;  
 DR TIGR; PSPT04792; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR002559; Transposase\_11.  
 DR Pfam; PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 486 AA; 55709 MW; 9690BE7816A32165 CRC64;

Query Match 47.2%; Score 50; DB 2; Length 486;  
 Best Local Similarity 58.8%; Pred. No. 26;  
 Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 3 IEGPLTROMLHGNGR 17  
 :|:|||||:  
 Db 431 VEHFPGNLKQWLFNGR 447

RESULT 9  
 0877Q7 PRELIMINARY; PRT; 486 AA.

AC 0877Q7;  
 DT 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
 DE ISPy6, transposase.  
 GN OrderedlocusNames=PSPT00358, PSPT03734;  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Davidse S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidse S.C., Brinkac L.M., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
 RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinour S., Chatterjee A.K.,  
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,  
 RA Bender C.L., White O., Fraser C.M., Collier A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 Pseudomonas syringae pv. tomato DC3000."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).

DR EMBL; AE016857; AA053902.1; -;  
 DR EMBL; AE016869; AA057203.1; -;  
 DR TIGR; PSPT00358; -;  
 DR TIGR; PSPT03734; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR002559; Transposase\_11.  
 DR Pfam; PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 486 AA; 55746 MW; 76343F2F8EBF7B2 CRC64;

Query Match 47.2%; Score 50; DB 2; Length 486;  
 Best Local Similarity 58.8%; Pred. No. 26;  
 Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 3 IEGPLTROMLHGNGR 17  
 :|:|||||:  
 Db 431 VEHFPGNLKQWLFNGR 447

RESULT 10  
 0877V4 PRELIMINARY; PRT; 486 AA.  
 AC 0877V4;  
 DT 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
 DE ISPy6, transposase.  
 GN OrderedlocusNames=PSPT01477, PSPT01567, PSPT01929, PSPT03808,  
 PSPT04060, PSPT04485;  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Davidse S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidse S.C., Brinkac L.M., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
 RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinour S., Chatterjee A.K.,  
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,



RA Bender C.L., White O., Fraser C.M., Collmer A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 RT Pseudomonas syringae pv. tomato DC3000."  
 RT Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).  
 DR EMBL; AE016861; AA054998.1; -  
 DR EMBL; AE016861; AA055087.1; -  
 DR EMBL; AE016862; AA055447.1; -  
 DR EMBL; AE016863; AA057277.1; -  
 DR EMBL; AE016870; AA057517.1; -  
 DR EMBL; AE016872; AA057934.1; -  
 DR TIGR; PSEP01477; -  
 DR TIGR; PSEP01567; -  
 DR TIGR; PSEP01929; -  
 DR TIGR; PSEP03808; -  
 DR TIGR; PSEP04060; -  
 DR TIGR; PSEP04485; -  
 DR GO; GO:0003677; P:DNA binding; IEA.  
 DR GO; GO:0004803; P:Transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR002559; Transposase\_11.  
 DR Pfam; PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 486 AA; 55718 MW; 0F90A6895854AD2 CRC64;  
 Query Match 47.2%; Score 50; DB 2; Length 486;  
 Best Local Similarity 58.8%; Pred. No. 26;  
 Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 QY 3 IEGP--TLRQWLHGNGR 17  
 DB 431 VEHFPGNTKQWLFNGNR 447  
 RESULT 11  
 ID EP42\_HUMAN STANDARD; PRT; 690 AA.  
 AC P16452;  
 DT 01-AUG-1990 (rel. 15, Created)  
 DT 01-AUG-1990 (rel. 15, Last sequence update)  
 DT 25-JAN-2005 (rel. 46, Last annotation update)  
 DE Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)  
 DE (P4.2)  
 GN Name=EP42; Synonyms=E42p;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC TISSUE=Reticulocytes;  
 RX MEDLINE=9121288; PubMed=2052563;  
 RA Korsgren C., Cohen C.M.;  
 RT "Organization of the gene for human erythrocyte membrane protein 4.2:  
 RT structural similarities with the gene for the a subunit of factor  
 RT XIII.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4840-4844 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (ISOFORM SHORT).  
 RC TISSUE=Reticulocytes;  
 RX MEDLINE=90138879; PubMed=2300550;  
 RA Korsgren C., Lawler U., Lambert S., Speicher D., Cohen C.M.;  
 RT "Complete amino acid sequence and homologues of human erythrocyte  
 RT membrane protein band 4.2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:613-617 (1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
 RC TISSUE=Reticulocytes;  
 RX MEDLINE=90138995; PubMed=1689063;  
 RA Sung L.A., Chien S., Chang L.-S., Lambert K., Bliss S.A.,  
 RA Boushesira E.E., Nagel R.L., Schwartz R.S., Rybicki A.C.;  
 RT "Molecular cloning of human protein 4.2: a major component of the  
 RT erythrocyte membrane.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:955-959 (1990).

RN [4]  
 RP MYRISTOYLATION.  
 RX MEDLINE=92184834; PubMed=1544941;  
 RA Ristinger M.A., Dolinas B.M., Cohen C.M.;  
 RT "Human erythrocyte protein 4.2, a high copy number membrane protein,  
 RT is N-myristylated.";  
 RL J. Biol. Chem. 267:5680-5685 (1992).  
 RN [5]  
 RP PHOSPHORYLATION SITE SER-247.  
 RX MEDLINE=93271204; PubMed=8499466; DOI=10.1016/0005-2736(93)90156-T;  
 RA Dolinas B., Speicher D.W., Gupta B., Cohen C.M.;  
 RT "Structural domain mapping and phosphorylation of human erythrocyte  
 RT pallidin (band 4.2).";  
 RL Biochim. Biophys. Acta 1148:19-29 (1993).  
 RN [6]  
 RP VARIANT HS THR-111.  
 RX MEDLINE=92216098; PubMed=1558976;  
 RA Boushesira E.E., Schwartz R.S., Yawata Y., Ata K., Kanzaki A.,  
 RA Oiu J.-H., Nagel R.L., Rybicki A.C.;  
 RT "An alanine-to-threonine substitution in protein 4.2 cDNA is  
 RT associated with a Japanese form of hereditary hemolytic anemia  
 RT (protein 4.2 Nippon).";  
 RL Blood 79:1846-1854 (1992).  
 RN [7]  
 RP VARIANT HS THR-111.  
 RX MEDLINE=95118828; PubMed=7819064;  
 RA Takeoka Y., Ideguchi H., Matsuda M., Sakamoto N., Takeuchi T.,  
 RA Fukumaki Y.;  
 RT "A novel mutation in the erythrocyte protein 4.2 gene of Japanese  
 RT patients with hereditary spherocytosis (protein 4.2 Fukuoka).";  
 RL Br. J. Haematol. 88:527-533 (1994).  
 RN [8]  
 RP VARIANT HS GLN-279.  
 RX MEDLINE=95290393; PubMed=7772513;  
 RA Hayette S., Morle L., Bozon M., Ghanem A., Ristinger M., Korsgren C.,  
 RA Tanner M.J.A., Fatoum S., Cohen C.M., Delaunay J.;  
 RT "A point mutation in the protein 4.2 gene (allele 4.2 Tozeur)  
 RT associated with hereditary haemolytic anaemia.";  
 RL Br. J. Haematol. 89:762-770 (1995).  
 CC -1- FUNCTION: Probably plays an important role in the regulation of  
 CC erythrocyte shape and mechanical properties.  
 CC -1- SUBUNIT: Oligomer. Interacts with the cytoplasmic domain of  
 CC SLC4A1/band 3 anion transport protein.  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (cytoplasmic surface of  
 CC erythrocyte membranes) and cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Short;  
 CC IsoId=P16452-1; Sequence=Displayed;  
 CC Note=Major isoform.  
 CC Name=Long;  
 CC IsoId=P16452-2; Sequence=VSP 006416;  
 CC -1- PTM: Both CAMP-dependent kinase (CAK) and another kinase present  
 CC in the red-blood cells seem to be able to phosphorylate EPB42.  
 CC -1- DISEASE: Defects in EPB42 are a cause of hereditary spherocytosis  
 CC (HS) [MIM:177070], a hematologic disorder leading to chronic  
 CC hemolytic anemia and characterized by numerous abnormally shaped  
 CC erythrocytes which are generally spheroidal. Absence of band 4.2  
 CC associated with spur or target erythrocytes has also been  
 CC reported.  
 CC -1- MISCELLANEOUS: The substitution of an Ala for a Cys in the active  
 CC site may be responsible for the lack of transglutaminase activity  
 CC of band 4.2.  
 CC -1- SIMILARITY: Belongs to the transglutaminase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR EMBL; M60298; AAA74589.1; -.
DR EMBL; L06519; AAA52385.1; -.
DR EMBL; L06447; AAA52385.1; JOINED.
DR EMBL; L06448; AAA52385.1; JOINED.
DR EMBL; L06449; AAA52385.1; JOINED.
DR EMBL; L06450; AAA52385.1; JOINED.
DR EMBL; L06511; AAA52385.1; JOINED.
DR EMBL; L06512; AAA52385.1; JOINED.
DR EMBL; L06513; AAA52385.1; JOINED.
DR EMBL; L06515; AAA52385.1; JOINED.
DR EMBL; L06516; AAA52385.1; JOINED.
DR EMBL; L06517; AAA52385.1; JOINED.
DR EMBL; M29339; AAA35798.1; -.
DR EMBL; M30646; AAA36402.1; -.
DR EMBL; M30647; AAA36401.1; -.
DR PIR; A39707; A39707.
DR HSSP; P52181; 1G0D.
DR Genew; HGNC:3381; EPB42.
DR MIM; 177070; -.
DR GO; GO:0005856; C:cytoskeleton; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0005524; F:ATP binding; TAS.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
DR InterPro; IPR001102; Gluttransfg.
DR InterPro; IPR008958; Transglut_C.
DR InterPro; IPR002931; Transglutase_like.
DR Pfam; PF00927; Transglut_C; 2.
DR Pfam; PF01841; Transglut_core; 1.
DR Pfam; PF00868; Transglut_N; 1.
DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.
KW Alternative splicing; Cell shape; Cytoskeleton;
KW Direct protein sequencing; Disease mutation; Erythrocyte maturation;
KW Hereditary hemolytic anemia; Lipoprotein; Myristate; Phosphorylation;
KW Structural protein.
FT SITE 0 0 By similarity.
FT SITE 30 38 Band 3 binding (By similarity).
FT LIPID 1 1 N-myristoyl glycine.
FT MOD_RES 247 247 Phosphoserine (by PKA) (Probable).
FT VARSPLIC 2 2 Q -> QGESQKSTGLAGLYAPAAAPVFKSGMD (in isoform long).
FT VARIANT 111 111 /FTId=VSP_006416.
FT VARIANT 111 111 A -> T (in HS; Nippon/Fukuoka).
FT VARIANT 279 279 R -> Q (in HS; Tozeur).
FT VARIANT 279 279 /FTId=VAR_012268.
FT CONFLICT 334 339 TRPALP -> KRGLPC (in Ref. 3).
FT CONFLICT 349 349 D -> H (in Ref. 3).
FT CONFLICT 375 375 V -> L (in Ref. 3).
SQ SEQUENCE 690 AA; 76841 MW; C6B605869A0A7A8B CRC64;

Query Match 47.2%; Score 50; DB 1; Length 690;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 PTLRQWLMHNGR 17
DB 249 PTLRQWLMHNGR 260

RESULT 12
EPB42_MOUSE STANDARD; PRT; 690 AA.
AC P49222;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)
DE (P4.2).
GN Name=Ep42; Synonyms=Ep44.2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Reticulocytes;
RA MEDLINE=95003352; PubMed=7919657;
RX Rybicki A.C., Schwartz R.S., Qiu J.U.-H., Gilman J.G.;
RT "Molecular cloning of mouse erythrocyte protein 4.2: a membrane
RT protein with strong homology with the transglutaminase supergene
RT family.";
RL Mamm. Genome 5:438-445(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; and C57BL/6J; TISSUE=liver, and Reticulocytes;
RX MEDLINE=95048323; PubMed=7959722;
RA Korgren C., Cohen C.M.;
RT "cDNA sequence, gene sequence, and properties of murine pallidin (band
RT 4.2), the protein implicated in the murine pallid mutation.";
RL Genomics 21:478-485(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Karscay B.B.K., Enzhong X.E.X., Chang L.-S.L.S.;
RT "Murine erythrocyte protein 4.2 gene: similarity and differences in
RT structure and expression from its human counterpart.";
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Probably plays an important role in the regulation of
CC erythrocyte shape and mechanical properties.
CC -! SUBUNIT: Oligomer. Interacts with the cytoplasmic domain of
CC SLC4A1/band 3 anion transport protein.
CC -! SUBCELLULAR LOCATION: Membrane-associated (cytoplasmic surface of
CC erythrocyte membranes) and cytoplasmic.
CC -! MISCELLANEOUS: The substitution of an Ala for a Cys in the active
CC site may be responsible for the lack of transglutaminase activity
CC of band 4.2.
CC -! SIMILARITY: belongs to the transglutaminase family.
CC -! CAUTION: was originally (Ref.2) thought to be pallidin.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03487; AAA62275.1; -.
DR EMBL; U04055; AAA67916.1; -.
DR EMBL; U04056; AAA67917.1; -.
DR EMBL; U35933; AAA39875.1; -.
DR PIR; A54741; A54741.
DR HSSP; 008188; 1L9M.
DR MGD; MGI:95402; Ep44.2.
DR InterPro; IPR001102; Gluttransfg.
DR InterPro; IPR008958; Transglut_C.
DR InterPro; IPR002931; Transglutase_like.
DR Pfam; PF00927; Transglut_C; 2.
DR Pfam; PF00868; Transglut_core; 1.
DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.
KW Cell shape; Cytoskeleton; Erythrocyte maturation; Lipoprotein;
KW Myristate; Phosphorylation; Structural protein.
FT SITE 0 0 By similarity.
FT SITE 30 38 Band 3 binding (By similarity).
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT MOD_RES 247 247 Phosphoserine (By similarity).
FT CONFLICT 21 21 Y -> H (in Ref. 2; AAA67917).
FT CONFLICT 223 223 K -> N (in Ref. 2; AAA67917).
FT CONFLICT 397 397 C -> S (in Ref. 2; AAA67917).
FT CONFLICT 449 449 K -> R (in Ref. 2; AAA67917).
FT CONFLICT 527 527 S -> R (in Ref. 2; AAA67917).
FT CONFLICT 620 620 S -> C (in Ref. 2 and 3).
SQ SEQUENCE 690 AA; 76608 MW; 3F6BCEP23DD385A6 CRC64;

```

Query Match 47.2%; Score 50; DB 2; Length 691;

RESULT 15			
MEAS_PENPA	STANDARD;	PRT;	1774 AA.
ID_MEAS_PENPA			
AC_P2367;			
DT_01-AUG-1991	(Rel. 19, Created)		
DT_01-AUG-1991	(Rel. 19, last sequence update)		
DT_05-JUN-2004	(Rel. 44, last annotation update)		
DE_6-methylsalicylic acid synthase (EC 2.3.1.165) (6-MSAS) .			
OS_Penicillium patulum (penicillium griseoflavum) .			
OC_Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.			
OX_NCBI_TaxID=5078;			
RN[1]			
RP_SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.			
RC_STRAIN=DSM 62862;			
RX_MEDLINE=91006137; PubMed=2209605;			
RA_Beck J, Ripka S, Siegener A, Schütz E, Schweizer E.;			
RT "The multifunctional 6-methylsalicylic acid synthase gene of			
RT penicillium patulum. Its gene structure relative to that of other			
RT polyketide synthases.";			
RL_Eur. J. Biochem. 192:487-498(1990) .			
CC -FUNCTION: This multifunctional enzyme is a polyketide synthase. It catalyzes a total of 11 steps by seven different component			

```

CC enzymes, in the biosynthesis of the antibiotic patulin.
CC -1- CARBOLYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH = 6-
CC methylsalicylate + 4 COA + 3 CO(2) + NADP(+).
CC -1- PATHWAY: Patulin biosynthesis.
CC -1- SUBUNIT: Homomultimer.
CC -1- INDUCTION: In the late logarithmic growth phase.
CC -1- SIMILARITY: Contains 1 acyl carrier domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X55776; CAA39295.1; -.
CC PIR; S13178; S13178.
CC InterPro; IPR001227; AC transferase.
CC InterPro; IPR009081; ACP-like.
CC InterPro; IPR000794; ketoacyl_synth.
CC InterPro; IPR006163; pp_bind.
CC InterPro; IPR006162; Ppantase.S.
CC Pfam; PF00698; Acyl transf 1; 1.
CC Pfam; PF00109; ketoacyl-synt; 1.
CC Pfam; PF02801; ketoacyl-synt_C; 1.
CC Pfam; PF00550; pp-binding; 1.
CC PROSITE; PS00075; ACP DOMAIN; 1.
CC PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
CC Antifibiotic biosynthesis; Direct protein sequencing;
CC Multifunctional enzyme; NADP; Phosphopantetheine; Transferase.
CC FT DOMAIN 186 238 Acyltransferase.
CC FT DOMAIN 642 676 Acetyl/malonyl transferases.
CC FT DOMAIN 1403 1450 2-oxoacyl reductase.
CC FT DOMAIN 1700 1769 Acyl carrier (ACP).
CC NP_BIND 1419 1424 NADP (Potential).
CC ACT_SITE 204 204 Beta-ketoacyl synthase (by similarity).
CC ACT_SITE 653 653 Malonyltransferase (by similarity).
CC BINDING 1732 1732 Phosphopantetheine (by similarity).
CC SEQUENCE 1774 AA; 190732 MW; 05ED5DD10863F938 CRC64;

Query Match 47.2%; Score 50; DB 1; Length 1774;
Best Local Similarity 44.4%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LAIEGPTLRQWLHGNGRD 18
|||:|||||:|
488 LALQAKTLDMWTAEKGD 505

Db 488 LALQAKTLDMWTAEKGD 505

RESULT 16
066272 PRELIMINARY; PRT; 245 AA.
AC 066272;
DT 01-NOV-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DR Photosynthetic reaction center L subunit (Fragment).
GN Name=pufl;
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAM14332;
RX MEDLINE=21822632; PubMed=1832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unspecified diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
EMBL; AB010981; BAA25791.1; -.

```

```

DR HSSP; P02954; 100V.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0045156; F:electron transporter; transferring electron. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis; light reaction; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR00484; Photo_RC.
DR Pfam; PF00124; Photo RC; 1.
DR PRINTS; PR00256; REACTCENTRE.
DR TIGRFAMs; TIGR0157; pufl; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
DR NON_TER 1
SQ SEQUENCE 245 AA; 27214 MW; 52B268713E199ABD CRC64;

Query Match 46.2%; Score 49; DB 2; Length 245;
Best Local Similarity 81.8%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AIEGPTLRQWL 12
|||||||
Db 25 AIEGPTLRQWL 35

RESULT 17
082989 PRELIMINARY; PRT; 249 AA.
AC 082989;
DT 01-NOV-1998 (TRENBLREL. 08, Created)
DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DR Photosynthetic reaction center L subunit (Fragment).
GN Name=pufl;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBIC3019;
RX MEDLINE=21822632; PubMed=1832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unspecified diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
EMBL; AB015708; BAA32995.1; -.
DR HSSP; P02954; 1YST.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0045156; F:electron transporter; transferring electron. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis; light reaction; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR00484; Photo_RC.
DR Pfam; PF00124; Photo RC; 1.
DR PRINTS; PR00256; REACTCENTRE.
DR TIGRFAMs; TIGR0157; pufl; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
DR NON_TER 1
SQ SEQUENCE 249 AA; 27702 MW; 4D68EBC82B7166AD CRC64;

Query Match 46.2%; Score 49; DB 2; Length 249;
Best Local Similarity 81.8%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AIEGPTLRQWL 12
|||||||
Db 25 AIEGPTLRQWL 35

RESULT 18
09XDVO PRELIMINARY; PRT; 278 AA.
AC 09XDVO;
DT 01-NOV-1999 (TRENBLREL. 12, Created)

```

DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE Photosynthetic reaction center L subunit.  
 GN Name=pslI;  
 OS Erythrobacter sp. MB13960.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;  
 OC Sphingomonadaceae; Erythrobacter.  
 OX NCBI\_TaxID=94771;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB13960;  
 RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;  
 RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,  
 RA Hamada T., Eisen J.A., Frazer C.M., Delong E.P.;  
 RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs";  
 RL Nature 415:630-633 (2002).  
 DR EMBL; AB027515; BAA78672.1; -.  
 DR HSSP; P02954; 1YST.  
 DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.  
 DR GO; GO:0045156; P:electron transporter, transferring electron. . .; IEA.  
 DR GO; GO:006118; P:electron transport; IEA.  
 DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.  
 DR InterPro; IPR005871; Photo\_L.  
 DR Pfam; PF00124; Photo\_RC; 1.  
 DR PRINTS; PR00256; REACTCENTRE.  
 DR TIGRPFAM; TIGR01157; pufL; 1.  
 DR PROSITE; PS00244; REACTION\_CENTER; 1.  
 SQ SEQUENCE 278 AA; 30735 MW; 0BE618844B3C54FB CRC64;

Query Match 46.2%; Score 49; DB 2; Length 278;  
 Best Local Similarity 81.8%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AIEGPTLRQWL 12  
 DB 54 AIEGPTLRPWL 64

## RESULT 19

ID 088D08 PRELIMINARY; PRT; 412 AA.

AC 088D08;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocNames=PP4765;  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Badnac L.M., Beaman M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,  
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,  
 RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,  
 RA Moazzes A., Uterback T.R., Rizzo W., Lee K., Kosack D., Noesl D.,  
 RA Medler H., Lauber J., Stjepandic D., Hohnsbeil J., Straetz M., Helm S.,  
 RA Kewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuenmler B.,  
 RA Frazer C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440";  
 RL Environ. Microbiol. 4:799-808 (2002).  
 DR EMBL; AB016791; AAN70335.1; -.  
 DR TIGR; PP4765; -.  
 DR GO; GO:0015036; P:disulfide oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000755; Adnrx\_redox.  
 DR InterPro; IPR001327; FAD\_dyt\_redox.

DR InterPro; IPR004792; HI0933 like.  
 DR InterPro; IPR000205; NAD\_BS-  
 DR InterPro; IPR001100; Pyr\_redox.  
 DR Pfam; PF03486; HI0933\_like; 1.  
 DR PRINTS; PR00419; ADKRDYASE.  
 DR PRINTS; PR00368; FADPNR.  
 DR PRINTS; PR00411; FNDRDYASRI.  
 DR ProDom; PD018041; HI0933 like; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 412 AA; 43682 MW; 1FAF5CE361F6D2C0 CRC64;

Query Match 46.2%; Score 49; DB 2; Length 412;  
 Best Local Similarity 50.0%; Pred. No. 31;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 EGPTRLQWLGNGRDT 19  
 DB 86 DADALRQWTHGIGRET 101

## RESULT 20

ID EREB\_ECOLI STANDARD; PRT; 419 AA.

AC P05789;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, last sequence update)  
 DT 05-JUL-2004 (Rel. 44, last annotation update)  
 DE Erythromycin esterase type II (EC 3.1.1.-).  
 GN Name=ereB;  
 OS Escherichia coli.  
 OC Plasmid pIP1527.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86259072; PubMed=3523438;  
 RA Arthur M., Autissier D., Courvalin P.;  
 RT "Analysis of the nucleotide sequence of the ereB gene encoding the  
 RT erythromycin esterase type II";  
 RL Nucleic Acids Res. 14:4987-4993 (1986).  
 CC -1- FUNCTION: This enzyme confers resistance to erythromycin through  
 CC inactivation by hydrolyzing the lactone ring of the antibiotic.  
 CC -1- MISCELLANEOUS: Erythromycin esterase type I and type II share no  
 CC significant homology except for the region from 279 to 309.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X03988; CAA7626.1; -.  
 DR PIR; A24381; EREBGM.  
 DR InterPro; IPR007815; Erythro\_esteras.  
 DR Pfam; PF05139; Erythro\_esteras; 1.  
 KW Antibiotic resistance; Hydrolase; Plasmid; Serine esterase.  
 SQ SEQUENCE 419 AA; 48173 MW; BCB07A565DBC8BA4 CRC64;

Query Match 46.2%; Score 49; DB 1; Length 419;  
 Best Local Similarity 46.7%; Pred. No. 32;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGPTRLQWLGNGRDT 18  
 DB 79 EGQJINMTHGQCTD 93

## RESULT 21

ID 07SF09 PRELIMINARY; PRT; 547 AA.

```

AC 07SFO9;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU09091.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OK NCBT_taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitchugh W., Ma L.-J., Smirnov S., Purcell S., Rehan B.,
RA Elkins T., Engels T., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selltremlkoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamel M., Kamywasell M., Mauceli E., Bielke C., Rudd S., Frislan D.,
RA Kytotova S., Raamsen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glas L., Orbach M.J., Berglund J., Voelker R.,
RA Varden O., Plaman M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Naryig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nuebaum C., Birren B.,
RT The Genome Sequence of the Filamentous Fungus Neurospora crassa.
RC Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AABX01000017; EAA35672.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR001395; Aldo/ket red.
DR InterPro; IPR002523; Mg2+ transportCoRa.
DR Pfam; PF01544; CoRa; 1.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 547 AA; 61554 MW; 41F0138ECFCF3E45 CRC64;

Query Match 46.2%; Score 49; DB 2; Length 547;
Best Local Similarity 72.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 PTLRQWLHNGR 16
DB 128 PTLRQWLHNGR 138

RESULT 22
Q6YS15 PRELIMINARY; PRT; 88 AA.
ID Q6YS15;
AC Q6YS15;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein B1446H1.10.
GN Name=B1446H1.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OK NCBT_taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RA Sasaki T., Matsumoto T., Karayose Y.;
DR EMBL; A006532; BAC99962.1; -
KW Hypothetical protein.
SQ SEQUENCE 88 AA; 8887 MW; F163E46B01AEF57 CRC64;

```

```

Query Match 45.8%; Score 48.5; DB 2; Length 88;
Best Local Similarity 76.9%; Pred. No. 7.3;
Matches 10; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 5 PTLRQWLHNGR 17
DB 16 PTLRQWLHNGR 27

RESULT 23
Q82S29 PRELIMINARY; PRT; 154 AA.
ID Q82S29;
AC Q82S29;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Glycoside hydrolase family 24 (EC 3.2.1.17).
GN OrderedLocusNames=NE2534;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OK NCBT_taxid=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RC MEDLINE=22586410; PubMed=12700255;
RX DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chaurin P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whitaker M.M., Atp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
heteropolymers of the prokaryotes cell walls.
CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 24 family.
CC EMBL; BX31865; CAD86446.1; -
DR GO; GO:0016798; F:hydrolase activity; acting on glycosyl bonds; IEA.
DR GO; GO:0003796; F:lysozyme activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0016935; P:cell wall catabolism; IEA.
DR GO; GO:0019835; P:cytolysis; IEA.
DR GO; GO:0042742; P:defense response to bacteria; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002196; Glyco_hydro_24.
DR Pfam; PF00959; Phage_lysozyme; 1.
KW Bacteriolytic enzyme; Complete proteome; Glycosidase; Hydrolase.
SQ SEQUENCE 154 AA; 17326 MW; BB51D426CA4BC89 CRC64;

Query Match 45.3%; Score 48; DB 2; Length 154;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLRQWLHNGR 17
DB 123 QGELRWVHGCGK 135

RESULT 24
Q9V492 PRELIMINARY; PRT; 168 AA.
ID Q9V492;
AC Q9V492;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE CG11077-PA (RE55125p).
GN ORFNames=CG11077;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

NCBI\_TaxId=7227;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celnikner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abriil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Houtin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshirei A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheier F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RT Science 287:2185-2195(2000).  
 [2]  
 SEQUENCE FROM N.A.  
 MEDLINE=2426065; PubMed=12537568;  
 RA Celnikner S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodegren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers B.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT melanogaster euchromatic genome sequence";  
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 SEQUENCE FROM N.A.  
 MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celnikner S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective";  
 RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 [4]  
 SEQUENCE FROM N.A.  
 MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker U.S., Millburn G.H., Prochownik K.S.,  
 RA Smith C.D., Tudy J.L., Whitfield E.J., Bayraktaroglu L., Bergman B.P.,  
 RA Beutenkourt B.R., Celnikner S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;

"Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review";  
 RT Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 [5]  
 SEQUENCE FROM N.A.  
 RA Flybase;  
 RT Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 [6]  
 SEQUENCE FROM N.A.  
 RA Flybase;  
 RT Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 [7]  
 SEQUENCE FROM N.A.  
 RA STRAIN-Berkeley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Munoz J., Nuncio J., Pacle J., Pargues V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celnikner S.;  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RA EMBL; AE003846; AAF59393.1; -;  
 DR EMBL; AY071482; AAL49104.1; -;  
 DR Flybase; FBgn003930; CG11077.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005778; C:peroxisomal membrane; IEA.  
 DR InterPro; IPR007248; MPv17\_PMP22.  
 DR Pfam; PF04117; MPv17\_PMP22; 1.  
 SQ SEQUENCE 168 AA; 19521 MW; 48E216A954E43D39 CRC64;

Query Match 45.3%; Score 48; DB 2; Length 168;  
 Best Local Similarity 61.5%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 1 LAIBPTLRQWLH 13  
 50 LVFVPTLRWYH 62

RESULT 25  
 ID 0688F0 PRELIMINARY; PRT; 392 AA.  
 AC 0688F0;  
 DT 25-OCT-2004 (TRMBLrel. 28, Created)  
 DT 25-OCT-2004 (TRMBLrel. 28, Last sequence update)  
 DE Hypothetical protein OSJNBa0035101.11.  
 GN Name=OSJNBa0035101.11;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxId=39947;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Lwu H.-L.,  
 RA Li Y.-F., Lin S.-Y., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.;  
 RT "Oryza sativa BAC OSJNBa0035101 genomic sequence";  
 RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC137000; AAU10721.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 392 AA; 42229 MW; 268075FEF48FC2F CRC64;

Query Match 45.3%; Score 48; DB 2; Length 392;  
 Best Local Similarity 58.8%; Pred. No. 42;  
 Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
 4 EGPT--LRQWLHNGRD 18  
 :|:|:|:|:|



Db 121 BGATRDSCQWLHGDD 137

RESULT 26

AC 067UT0 PRELIMINARY; PRT; 396 AA.

DT 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Putative serine/threonine protein kinase.

GN Name=P0046G12.38;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

OC NCB1\_TaxID=39947;

OX [1]

RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nippondare (GA3) genomic DNA, chromosome 9, PAC clone:P0046G12.";

RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF005419; BAD38089.1; -

DR GO; GO:0016301; F.kinase activity; IEA.

DR CO; GO:0004674; F.protein serine/threonine kinase activity; IEA.

DR InterPro; IPR011009; Kinase\_like.

DR InterPro; IPR000719; Prot\_Kinase.

DR InterPro; IPR002290; Ser\_thr\_pkinase.

DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.

DR InterPro; IPR001245; Tyr\_pkinase.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00220; S\_TKC; 1.

DR SMART; SM00219; TyKc; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

KW Kinase; Serine/threonine-protein kinase.

SW SEQUENCE 396 AA; 43831 MW; B2C9FE83CF344456 CRC64;

Query Match 45.3%; Score 48; DB 2; Length 396;

Best local Similarity 66.7%; Pred. No. 43;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 IEGLTLRQWLHG 14

Db 193 LTGATLQWLHG 204

RESULT 27

AC 06LP74 PRELIMINARY; PRT; 631 AA.

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein BL2800.

GN Name=BL2800; OrderedLocusNames=BPRA2520;

OS Photobacterium profundum (Photobacterium sp. (strain SS9)).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Photobacterium.

OC NCB1\_TaxID=74109;

OX [1]

RP SEQUENCE FROM N.A.

RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F., Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D., Valle G.;

RT "Genome analysis of Photobacterium profundum reveals the complexity of high pressure adaptations.";

RT Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; CR378671; CAG20902.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR011009; Kinase\_like.

DR InterPro; IPR001932; PP2C-like.

DR InterPro; IPR000719; Prot\_kinase.

DR InterPro; IPR002290; Ser\_thr\_pkinase.

DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.

DR InterPro; IPR001245; Tyr\_pkinase.

DR Pfam; PF00069; Pkinase; 1.

DR Pfam; PF00481; PP2C; 1.

DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00332; PP2C; 1.

DR SMART; SM00331; PP2C\_SIG; 1.

DR SMART; SM00220; S\_TKC; 1.

DR SMART; SM00219; TyKc; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.

KW Complete proteome; Hypothetical protein.

SW SEQUENCE 631 AA; 71260 MW; 4AC862AC45B58C06 CRC64;

Query Match 45.3%; Score 48; DB 2; Length 631;

Best local Similarity 53.8%; Pred. No. 70;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 IEGLTLRQWLHG 15

Db 404 VEGSLRQWMDN 416

RESULT 28

EP42 BOVIN STANDARD; PRT; 686 AA.

AC 0465T0

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)

DE Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2) (P4.2)

GN Name=EPB42; Synonyms=BBP42;

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

OC NCB1\_TaxID=9913;

OX [1]

RP SEQUENCE FROM N.A.

RA STRAIN=Japanese black;

RX MEDLINE=98244826; PubMed=9576866;

RA Matsumoto M., Inaba M., Ono K.-I.;

RT "Molecular basis of bovine red-cell protein 4.2 polymorphism in Japanese black cattle.";

RL Biochem. J. 332:183-187(1998).

CC -1- FUNCTION: Band 4.2 probably plays an important role in the regulation of erythrocyte shape and mechanical properties. The major membrane binding for band 4.2 is the cytoplasmic domain of the erythrocyte anion transporter, band 3 (By similarity).

CC -1- SUBUNIT: Oligomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic surface of erythrocyte membranes (By similarity).

CC -1- MISCELLANEOUS: The substitution of an Ala for a Cys in the active site may be responsible for the lack of transglutaminase activity of band 4.2.

CC -1- SIMILARITY: Belongs to the transglutaminase family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib.ch](mailto:license@sib.ch)).



CC -----  
 DR EMBL; AF030030; AAC48855.1; -.  
 DR HSSP; 008188; 1L9M.  
 DR InterPro; IPR001102; GlutransfG.  
 DR InterPro; IPR008958; Transglut C.  
 DR InterPro; IPR002931; Transglutase\_like.  
 DR Pfam; PF00927; Transglut\_C; 2.  
 DR Pfam; PF01841; Transglut\_core; 1.  
 DR Pfam; PF00868; Transglut\_N; 1.  
 DR SMART; SM00460; TGC; 1.  
 DR PROSITE; PS00547; TRANSGLUTAMINASES; FALSE NEG.  
 DR Cell shape; Cytokeleton; Erythrocyte maturation; Lipoprotein;  
 KW Myristate; Phosphorylation; Structural protein.  
 FT INIT\_MET 0  
 FT SITE 30 38 By similarity.  
 FT LIPID 1 1 N-myristoyl glycine (By similarity).  
 FT MOD\_RES 246 246 Phosphoserine (By similarity).  
 FT SO SEQUENCE 686 AA; 76485 MW; 71CDB6CC82FE7D CRC64;

Query Match 45.3%; Score 48; DB 1; Length 686;  
 Best Local Similarity 66.7%; Pred. No. 76;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 PTLROWLHNGNR 17  
 DB 248 PTLROWVTGHR 259

RESULT 29  
 ID 046509 PRELIMINARY; PRT; 687 AA.  
 AC 046509;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Erythrocyte protein 4.2.  
 GN Name=BBP42;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98244826; PubMed=9576866;  
 RA Matsumoto M., Inaba M., Ono K.;  
 RT "Molecular basis of bovine red-cell protein 4.2 polymorphism in  
 RT Japanese black cattle.";  
 RL Biochem. J. 332:183-187(1998).  
 DR EMBL; AF030029; AAC48854.1; -.  
 DR HSSP; 008188; 1L9M.  
 DR GO; GO:0018149; P:peptide cross-linking; IEA.  
 DR InterPro; IPR001102; GlutransfG.  
 DR InterPro; IPR002114; HPr\_Serp\_S.  
 DR InterPro; IPR008958; Transglut\_C.  
 DR InterPro; IPR002931; Transglutase\_like.  
 DR Pfam; PF00927; Transglut\_C; 2.  
 DR Pfam; PF01841; Transglut\_core; 1.  
 DR Pfam; PF00868; Transglut\_N; 1.  
 DR SMART; SM00460; TGC; 1.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
 DR SEQUENCE 687 AA; 76616 MW; 599708686B355D2D CRC64;

Query Match 45.3%; Score 48; DB 2; Length 687;  
 Best Local Similarity 66.7%; Pred. No. 76;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 PTLROWLHNGNR 17  
 DB 249 PTLROWVTGHR 260

RESULT 30

084FF9  
 ID 084FF9 PRELIMINARY; PRT; 1916 AA.  
 AC 084FF9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative hemagglutinin.  
 GN Name=agmC;  
 OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; Delta proteobacteria; Myxococcales;  
 OC Cyrobacterineae; Myxococcaceae; Myxococcus.  
 NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hartzell P.L., Youderian P.A.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY204461; AA022852.1; -.  
 DR InterPro; IPR003961; FN\_III.  
 DR SMART; SM00060; FN3; 2.  
 SO SEQUENCE 1916 AA; 199157 MW; 7BEA8D84ABD8A94B CRC64;

Query Match 45.3%; Score 48; DB 2; Length 1916;  
 Best Local Similarity 55.6%; Pred. No. 2.2e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LAIEGPTLRQWLHNGRD 18  
 DB 1483 LAEGHTLRVWARGGRE 1500

RESULT 31  
 ID 075174 PRELIMINARY; PRT; 182 AA.  
 AC 075174;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein OSUJBa0083F15.8.  
 GN Name=OSUJBa0083F15.8;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Buell C.R., Yuan Q., Gonyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.U., Jin S.S.,  
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Uterback T.T., Feldlyum T.V.,  
 RA Yang Q.Q., Haas B.U., Sun B.B., Peterson J.U., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Buell R.; (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC133398; AAR01777.1; -.  
 DR Hypothetical protein.  
 SO SEQUENCE 182 AA; 19525 MW; 12E4793265F0759B CRC64;

Query Match 44.3%; Score 47; DB 2; Length 182;  
 Best Local Similarity 61.5%; Pred. No. 27;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 TLRQWLHNGRD 19  
 DB 70 TLRPWLHNGDANT 82

RESULT 32  
 ID 055838 PRELIMINARY; PRT; 344 AA.  
 AC 055838

```

AC 055838;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NS5 protein (Fragment).
GN Name=NS5;
OS Yokose virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=64294;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-Oita 36;
RX MEDLINE=98080391; PubMed=9420202;
RA Kuno G., Chang G.J., Tsuchiya K.R., Karabatsos N., Cropp C.B.;
RT "Phylogeny of the genus Flavivirus."
RL J. Virol. 72:73-83 (1998).
DR EMBL; AF013414; AAC58802.1; -;
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR00208; Flavi_NS5.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF00972; Flavi_NS5.1.
FT NON_TER 1 1
FT NON_TER 344 344
SQ SEQUENCE 344 AA; 39610 MW; 3E791624EBCDC6D CRC64;

Query Match 44.3%; Score 47; DB 2; Length 344;
Best Local Similarity 64.3%; Pred. No. 53;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 GPTLRQMLHNGNRD 18
DB 182 GNTLMQMLNENGED 195

RESULT 33
MOA CAUCR STANDARD; PRT; 349 AA.
ID MOA CAUCR
AC 09AC38;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Molybdenum cofactor biosynthesis protein A.
GN Name=moaA; OrderedLocustNames=CC0018;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=15892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=2113668; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nieman W.C., Feldberg T.V., Iaub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA DeBoy R.T., Dodson R.J., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolonay J.F., Smit J., Craven M.B., Knouri H.M., Shetty J.,
RA Berry K.J., Utecht T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
OC Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -1- FUNCTION: Together with moaC, is involved in the conversion of a
CC guanosine derivative (GMP) into molybdopterin precursor Z (By
CC similarity).
CC -1- COFACTOR: Binds 1 3Fe-4S cluster (By similarity).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis; first step.
CC -1- SIMILARITY: Belongs to the radical SAM superfamily. MoaA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005676; AKK22006.1; -;
DR PIR; B87251; B87251.
DR TIGR; CC0018; -;
DR HAMAP; MF_01225; -; 1.
DR InterPro; IPR006638; E1p3/M1aB/M1fB.
DR InterPro; IPR000385; MoaA_M1fB_PqGE.
DR InterPro; IPR010505; Mob_Synth_C.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF06463; Mob_Synth_C; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; E1p3; 1.
DR PROSITE; PS01305; MOA_N1fB_POGE; 1.
DR 3Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
KW Molybdenum cofactor biosynthesis.
FT METAL 42 42 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 49 49 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 275 275 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 292 292 Iron-sulfur (3Fe-4S) (Potential).
SQ SEQUENCE 349 AA; 37577 MW; 2D2E7C5273F7A170 CRC64;

Query Match 44.3%; Score 47; DB 1; Length 349;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGPTLRQMLHNGNRD 18
DB 191 EIPALIQMAHNGRCD 205

RESULT 34
O9HX67 PRELIMINARY; PRT; 391 AA.
ID O9HX67;
AC 09HX67;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PA3949;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=2043737; PubMed=10984043; DOI=10.1038/35023079;
RX Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Rolger K.R., Kes A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
OC Nature 406:959-964 (2000).
RL EMBL; AF004813; AAC07336.1; -;
DR PIR; E83151; E83151.
DR InterPro; IPR004792; H10933_1like.
DR Pfam; PF03486; H10933_1like.
DR ProDom; PD018041; H10933_1like; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 391 AA; 42423 MW; 46F4B0A76BAEC200 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 391;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 4 EGFTRQWHLHGNGRDT 19
DB 64 DADALRAWIHGIDT 79

RESULT 35
Q8S6J0 PRELIMINARY; PRT; 483 AA.
AC Q8S6J0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical protein OSUNB0023M11.3.
GN Name=OSUNB0023M11.3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatroidae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Niponbare;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nacimento L., Baker J., Santos L., Zutavern T., Miller B.,
RA Cumulus D.M., Katzenberger F., Muller S., Bell M., Balija V., Shah R.,
RA King L., Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC092749; AM08549.1; -.
DR Gramene; Q8S6J0; -.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR005162; Retrotrans_gag.
DR InterPro; IPR008916; Retrov_capsid_C.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR Hypothetical protein.
KM SEQUENCE 483 AA; 54631 MW; 1EAC123961BF76F6 CRC64;
SQ

Query Match 44.3%; Score 47; DB 2; Length 483;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 TLKQWHLHGNGRDT 19
DB 12 SVRSMHLGPRDT 24

RESULT 36
Q7G676 PRELIMINARY; PRT; 483 AA.
AC Q7G676;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein OSUNB0078C13.19 (putative gag-pol).
GN Name=OSUNB0078C13.19; ORFNames=OSUNB0023M11.3;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatroidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nacimento L., Zutavern T., Balija V., Bell M., Baker J.,
RA Miller B., Katzenberger F., Muller S., King L., Sullivan P., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RA "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569 (2003).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC123594; AAM74436.1; -.
DR EMBL; AC123594; AAM74436.1; -.
DR EMBL; AE017081; AAP53204.1; -.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR005162; Retrotrans_gag.
DR InterPro; IPR008916; Retrov_capsid_C.
DR Pfam; PF03732; Retrotrans_gag; 1.
KM Hypothetical protein.
SQ SEQUENCE 483 AA; 54631 MW; 1EAC123961BF76F6 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 483;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 TLKQWHLHGNGRDT 19
DB 12 SVRSMHLGPRDT 24

RESULT 37
Q7T918 PRELIMINARY; PRT; 3425 AA.
AC Q7T918;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Yokose virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=64294;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Oita 36;
RA Tajima S., Takasaki T., Yabe S., Matsuno S., Nomura H., Kurane I.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB114858; BAC79364.1; -.
DR HSSP; Q9Q4T1; IBER.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR01122; Flavi_capsid_C.
DR InterPro; IPR011492; Flavi_DEAD.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR001157; Flavi_NS1.
DR InterPro; IPR000752; Flavi_NS2A.
DR InterPro; IPR000487; Flavi_NS2B.
DR InterPro; IPR000404; Flavi_NS4A.
DR InterPro; IPR001528; Flavi_NS4B.
DR InterPro; IPR000208; Flavi_NS5.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001850; Peptidase_S7.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR002877; RnmJfcd_mtfase.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF07652; Flavi_DEAD; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF02832; Flavi_glycop_C; 1.

```

```

DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF00948; Flavi_NSI; 1.
DR Pfam: PF01005; Flavi_NSI2; 1.
DR Pfam: PF01002; Flavi_NSI2B; 1.
DR Pfam: PF01350; Flavi_NSI4; 1.
DR Pfam: PF01349; Flavi_NSI4B; 1.
DR Pfam: PF00972; Flavi_NSI5; 1.
DR Pfam: PF01570; Flavi_NSI5B; 1.
DR Pfam: PF01728; Flavi_NSI5C; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00949; Helicase_S7; 1.
DR Pfam: PD001556; Flavi_Glycoprote; 1.
DR Pfam: PD001496; Flavi_NSI1; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC_C; 1.
DR ATP-binding; Helicase; Hydrolase; Polypeptide.
KW SEQUENCE 3425 AA; 384202 MW; 467B9A5CEA1C1EB CRC64;
SQ

Query Match 44.3%; Score 47; DB 2; Length 3425;
Best Local Similarity 64.3%; Pred. No. 6e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 GPTLRQWLHGNGRD 18
Db 3163 GNTLQWLQMLNENGRD 3176

RESULT 38
Q7MP20 PRELIMINARY; PRT; 495 AA.
ID Q7MP20;
AC Q7MP20;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein VV0544.
GN OrderedLocustNames=VV0544;
OS Vibrio vulnificus (strain VV016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibriio.
OC NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Shang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-W., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
pathogen.";
RT Genome Res. 13:2577-2587(2003).
DR EMBL: AP005332; BAC93308.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 57275 MW; 2C121F9AC6051DB4 CRC64;

Query Match 43.9%; Score 46.5; DB 2; Length 495;
Best Local Similarity 52.6%; Pred. No. 94;
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 2 AIEG-----PTLRQWLHNG 15
Db 99 AMEGSSRVLPPTLAAMHLAN 117

RESULT 39
Q931U2 PRELIMINARY; PRT; 600 AA.
ID Q931U2;
AC Q931U2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BldKB, putative ABC transport system Iipoprotein.
GN ORFNames=SCBAC31E11.09;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

```

```

OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiese R., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RT Nature 417:141-147(2002).
CC -1- SIMILARITY: Belongs to the bacterial extracellular solute-binding
protein family 5.
DR EMBL: AL939122; CAC44320.1; -.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
KW Complete proteome; Iipoprotein.
SQ SEQUENCE 600 AA; 65533 MW; 9CAD7DF4B5D95E95 CRC64;

Query Match 43.9%; Score 46.5; DB 2; Length 600;
Best Local Similarity 64.3%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 4 EGPT-LRQWLHGNG 16
Db 174 DGPTVLRQWLHGNG 187

RESULT 40
P72407 PRELIMINARY; PRT; 602 AA.
ID P72407;
AC P72407;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BldKB.
GN Name=BldKB;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=M145;
RA Nowell J. R., McGovern K., Josick R.;
RT Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the bacterial extracellular solute-binding
protein family 5.
DR EMBL: U68036; AAB09555.1; -.
DR PIR: T45278; T45278.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
KW Complete proteome.
SQ SEQUENCE 602 AA; 65851 MW; 5CTF74FC4C9C0FBA CRC64;

Query Match 43.9%; Score 46.5; DB 2; Length 602;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 4 EGPT-LRQWLHGNG 16
Db 174 DGPTVLRQWLHGNG 187

```

```

RESULT 41
Q8RSL6 PRELIMINARY; PRT; 85 AA.
ID Q8RSL6
AC Q8RSL6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical protein.
OC Bacteroides bacterium.
OS uncultured bacterium.
OG Plasmid pB4.
OC Bacteria; environmental samples.
NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22477408; PubMed=12589432;
RA Tsuchi A., Schluter A., Bischoff N., Goesmann A., Meyer F., Puhler A.;
RT "The 79,370-bp conjugative plasmid pB4 consists of an IncP- $\beta$ 
RT backbone loaded with a chromate resistance transposon, the strA-strB
RT streptomycin resistance gene pair, the oxacillinase gene blaNDP-1, and
RT a tripartite antibiotic efflux system of the resistance-nodulation-
RT division family."
RL Mol. Genet. Genomics 268:570-584(2003).
DR EMBL: AJ231260; CAD24344.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 85 AA; 9445 MW; DC06B2B0121D5EB4 CRC64;

Query Match 43.4%; Score 46; DB 2; Length 85;
Best Local Similarity 47.4%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LAIEPTLROWLHNGRDT 19
Db 50 MEIEASCLAEAMLRGHFRDT 68

RESULT 42
Q9NB63 PRELIMINARY; PRT; 256 AA.
ID Q9NB63
AC Q9NB63;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Galactose-binding protein.
OS Tachypleus tridentatus (Japanese horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Tachypleus.
NCBI_TaxID=6853;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167756; PubMed=11133989; DOI=10.1074/jbc.M008414200;
RA Chen S.-C., Yeh C.-H., Yeh M.-S., Huang C.-J., Liu T.-Y.;
RT "Biochemical properties and cDNA cloning of two new lectins from the
RT plasma of Tachypleus tridentatus. Tachypleus plasma lectin 1 and 2."
RL J. Biol. Chem. 276:9631-9639(2001).
DR EMBL: AF264067; AAF74773.1; -.
DR InterPro: IPR006624; TCCPR.
DR SMART: SM00706; TCCPR; 6.
SQ SEQUENCE 256 AA; 28517 MW; SEC0272B88F44F8 CRC64;

Query Match 43.4%; Score 46; DB 2; Length 256;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 IEGLPTLROWLHNGR 17
Db 18 VVSPTLAEWTHINCK 32

RESULT 43
Q6HSY4 PRELIMINARY; PRT; 279 AA.
ID Q6HSY4
AC Q6HSY4;

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative receptor-like kinase.
GN Name=P0620H05.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39477;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: AP005394; BAD25865.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002230; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PSS0108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 279 AA; 31269 MW; F201A2700EF5F2EF CRC64;

Query Match 43.4%; Score 46; DB 2; Length 279;
Best Local Similarity 61.5%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 GPTLROWLHNGR 17
Db 85 GCCLHEWLNRR 97

RESULT 44
MOA_RHIME STANDARD; PRT; 349 AA.
ID MOA_RHIME
AC Q92PB4;
DT 26-FEB-2003 (Rel. 41, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Molybdenum cofactor biosynthesis protein A.
GN Name=moaA; OrderedAccession=U01864; ORFName=SMC00144;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kles E., Lelaire V., Masny D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weisner S., Gallibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC 1- FUNCTION: Together with moaC, is involved in the conversion of a
CC guanosine derivative (GMP) into molybdopterin precursor Z (By

```

```

CC similarity).
CC -1- COFACTOR: Binds 1 3Fe-4S cluster (By similarity).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis, first step.
CC -1- SIMILARITY: Belongs to the radical SAM superfamily. MoaA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL591788; CAC46443.1; -.
DR HAMAP: MF_01225; -.
DR InterPro: IPR006638; EIP3/MiaB/NiFe.
DR InterPro: IPR000385; MoaA_NiFe_PqGE.
DR InterPro: IPR010505; Mob synth C.
DR InterPro: IPR007197; Radical SAM.
DR Pfam: PF06463; Mob synth C; I.
DR Pfam: PF04055; Radical SAM; I.
DR SMART: SM00729; EIP3; I.
DR PROSITE: PS01305; MOA_NiFe_PQGE; 1.
DR 3Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
DR Molybdenum cofactor biosynthesis.
KW METAL
FT METAL 40 40 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 47 47 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 273 273 Iron-sulfur (3Fe-4S) (Potential).
FT METAL 290 290 Iron-sulfur (3Fe-4S) (Potential).
SQ SEQUENCE 349 AA; 38915 MW; 768B3A86EFD1C0A6 CRC64;

Query Match 43.4%; Score 46; DB 1; Length 349;
Best Local Similarity 53.3%; Pred. NO. 78;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 EGGPTLRQWLHNGNRD 18
DB 189 EIEELMRMAHGRGMD 203

RESULT 45
ID 074061 PRELIMINARY; PRT; 434 AA.
AC 074061;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glutamate-1-semialdehyde aminotransferase.
GN Name=glat;
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Thermoprotei; Cenarchaeales; Cenarchaeaceae;
OC Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B.
RX MEDLINE=8422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
RL J. Bacteriol. 180:5003-5009(1998).
CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
CC EMBL: AF083072; AAC62704.1; -.
CC PIR: T31313; T31313.
DR HSSP: P24630; 4GSA.
DR GO: GO:0030170; F:pyridoxal phosphate binding; IEA.
DR GO: GO:0008483; F:transaminase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro: IPR005814; Aminotrans_3.
DR Pfam: PF00202; Aminotran 3; 1.
DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; UNKNOWN_1.

```

```

KW Aminotransferase; Pyridoxal phosphate; Transferase.
SQ SEQUENCE 434 AA; 46931 MW; 171C80F06B3C6F25 CRC64;

Query Match 43.4%; Score 46; DB 2; Length 434;
Best Local Similarity 41.2%; Pred. NO. 98;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIEGPTLRQWLHNGNRD 18
DB 81 AVEGQLRGMWHTANE 97

```

Search completed: September 1, 2005, 16:21:14  
Job time : 74.6691 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 64.3597 Seconds

(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-12

Perfect score: 85

Sequence: 1 CADGPTLRWISFC 14

## Scoring table:

BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

## Database :

A\_Geneseq\_16Dec04:\*  
1: geneseqp19808:\*  
2: geneseqp19908:\*  
3: geneseqp20008:\*  
4: geneseqp20018:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20048:\*  
8: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	14	2	AAW09466 Thrombopo
2	85	100.0	14	2	AAW09462 Thrombopo
3	85	100.0	14	2	AAW09465 Thrombopo
4	85	100.0	14	2	AAW09482 Thrombopo
5	85	100.0	14	2	AAW33031 Thrombopo
6	85	100.0	14	2	AAW36633 Thrombopo
7	85	100.0	14	2	AAW33029 Thrombopo
8	85	100.0	14	2	AAW35401 Thrombopo
9	85	100.0	14	2	AAW36647 Thrombopo
10	85	100.0	14	2	AAW35400 Thrombopo
11	85	100.0	14	2	AAW33032 Thrombopo
12	85	100.0	14	3	AAW17014 Thrombopo
13	85	100.0	14	4	AAW25826 Human thr
14	85	100.0	14	4	AAU25852 Human thr
15	85	100.0	14	4	AAU25866 Human thr
16	85	100.0	14	5	ABW72900 TPO mimet
17	85	100.0	14	7	ADJ73051 TPO mimet
18	85	100.0	14	8	ADJ52686 CHI delet
19	85	100.0	14	8	ADJ51647 CHI delet
20	85	100.0	18	2	AAW09456 Thrombopo
21	85	100.0	18	2	AAW33023 Thrombopo
22	85	100.0	18	3	AAW17020 TPO-mimet
23	85	100.0	18	4	AAU25820 Human thr
24	85	100.0	18	5	ABW72906 TPO mimet
25	85	100.0	18	7	ADJ73058 TPO mimet

26	85	100.0	18	8	ADJ52693 CHI delet
27	85	100.0	18	8	ADJ51654 CHI delet
28	85	100.0	19	2	AAW09458 Thrombopo
29	85	100.0	19	2	AAW33025 Thrombopo
30	85	100.0	19	2	AAW25822 Human thr
31	76	89.4	13	2	AAW09467 Thrombopo
32	76	89.4	13	2	AAW35399 Thrombopo
33	76	89.4	13	2	AAW35417 Thrombopo
34	76	89.4	13	2	AAW33033 Thrombopo
35	76	89.4	13	2	AAW35413 Thrombopo
36	76	89.4	13	2	AAW35406 Thrombopo
37	76	89.4	13	2	AAW35422 Thrombopo
38	76	89.4	13	2	AAW35397 Thrombopo
39	76	89.4	13	4	AAU25997 Human thr
40	76	89.4	14	2	AAW35398 Thrombopo
41	76	89.4	14	2	AAW35386 Thrombopo
42	76	89.4	14	2	AAW35402 Thrombopo
43	76	89.4	14	2	AAW35404 Thrombopo
44	76	89.4	14	4	AAU25987 Human thr
45	76	89.4	14	4	AAU25983 Human thr
46	76	89.4	14	4	AAU25985 Human thr
47	72	84.7	12	2	AAW35423 Thrombopo
48	72	84.7	12	2	AAU26000 Human thr
49	67	78.8	13	2	AAW35404 Thrombopo
50	67	78.8	13	2	AAW35405 Thrombopo
51	67	78.8	13	4	AAU25994 Human thr
52	67	78.8	13	4	AAU25991 Human thr
53	67	78.8	13	4	AAU25990 Human thr
54	67	78.8	14	2	AAW35412 Thrombopo
55	67	78.8	14	2	AAW35407 Thrombopo
56	67	78.8	14	2	AAW35408 Thrombopo
57	67	78.8	14	2	AAW35403 Thrombopo
58	67	78.8	14	4	AAU25989 Human thr
59	67	78.8	14	4	AAU25989 Human thr
60	67	78.8	14	4	AAU25995 Human thr
61	67	78.8	14	4	AAU25992 Human thr
62	67	78.8	14	4	AAU25986 Human thr
63	67	78.8	14	4	AAU25988 Human thr
64	67	78.8	25	4	AAU26042 Human thr
65	67	78.8	25	8	ADM72531 TPO mimet
66	66	77.6	11	2	AAW35425 Thrombopo
67	66	77.6	11	4	AAU26001 Human thr
68	65	76.5	13	4	AAU26041 Human thr
69	64	75.3	14	3	AAW17017 TPO-mimet
70	64	75.3	14	5	ABW72903 CHI delet
71	64	75.3	14	8	ADJ51650 Human thr
72	64	75.3	10	2	AAW35427 Thrombopo
73	60	70.6	10	4	AAU26002 Human thr
74	60	70.6	12	8	AAU26039 Human thr
75	57	67.1	13	4	AAU26039 Human thr
76	57	67.1	13	8	ADM72529 TPO mimet
77	57	67.1	13	8	ADM72528 TPO mimet
78	57	67.1	14	2	AAW66732 Peptide C
79	57	67.1	14	4	AAU26040 Human thr
80	57	67.1	14	3	AAW17015 TPO-mimet
81	56	65.9	13	5	ABW72901 TPO mimet
82	56	65.9	13	7	ADJ73054 TPO mimet
83	56	65.9	13	7	ADJ73052 TPO mimet
84	56	65.9	13	7	ADJ73056 TPO mimet
85	56	65.9	13	7	ADJ73053 TPO mimet
86	56	65.9	13	7	ADJ73055 TPO mimet
87	56	65.9	13	7	ADJ73055 TPO mimet
88	56	65.9	13	8	ADJ52687 CHI delet
89	56	65.9	13	8	ADJ51648 CHI delet
90	55	64.7	15	3	AAW17018 TPO-mimet
91	55	64.7	15	5	ABW72904 TPO mimet
92	55	64.7	15	8	ADJ52691 CHI delet
93	55	64.7	15	8	ADJ52690 CHI delet
94	55	64.7	15	8	ADJ51652 CHI delet
95	55	64.7	15	8	ADJ51651 CHI delet
96	55	64.7	18	7	ADN59672 Thrombopo
97	55	64.7	22	7	ADN59839 TWP pepti
98	55	64.7	25	7	ADN59744 Thrombopo

99 54 63.5 14 2 AAW09479 Aaw09479 Thrombopo  
100 54 63.5 14 2 AAW36630 Aaw36630 Thrombopo

## ALIGNMENTS

RESULT 1  
ID AAW09466 standard; protein; 14 AA.  
XX AAW09466;  
AC AAW09466;  
XX  
DT 10-SEP-1997 (first entry)  
XX  
DE Thrombopoietin receptor binding compound cyclic peptide.  
XX  
KM Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;  
KW bone marrow transfusion; chemotherapy; radiation therapy.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..14  
FT Modified-site /note= "in acetyl form"  
FT Modified-site 14  
FT /note= "in amide form"  
XX  
PN WO640189-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 05-JUN-1996; 96WO-US008998.  
XX  
PR 07-JUN-1995; 95US-00472371.  
PR 07-JUN-1995; 95US-00473604.  
PR 07-JUN-1995; 95US-00476168.  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00484090.  
PR 07-JUN-1995; 95US-00485301.  
XX  
PA (GLAXO) GLAXO GROUP LTD.  
PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX  
DR WPI; 1997-051883/05.  
XX  
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide  
PT mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.  
XX  
PS Claim 30; Page 91; 106pp; English.  
XX  
CC The present sequence is a compound which binds to thrombopoietin (TPO)  
CC receptor (TR). The compound can be used for treating patients suffering  
CC from haematological disorders and thrombocytopenia resulting from  
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide  
CC may also be used to maintain the proliferation and growth of TPO-  
CC dependent cell lines and for use in biological research, for detecting  
CC TPO receptors on living cells  
XX  
SQ Sequence 14 AA;  
XX  
Query Match 100.0%; Score 85; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. NO. 6.9e-07; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CADGPTLRWISFC 14  
DB 1 CADGPTLRWISFC 14

RESULT 2  
ID AAW09462 standard; protein; 14 AA.  
XX AAW09462;  
AC AAW09462;  
XX  
DT 10-SEP-1997 (first entry)  
XX  
DE Thrombopoietin receptor binding compound peptide.  
XX  
KM Haematology; thrombocytopenia; TPO; TR; proliferation;  
KW bone marrow transfusion; chemotherapy; radiation therapy.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1..14  
FT /note= "Preferably linkages are selected from: -  
FT CH2OC(O)NR-; Phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6  
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is  
FT lower alkyl"  
FT Modified-site 1  
FT /note= "Preferably N-terminus is selected from: -NRR1; -  
FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;  
FT benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3  
FT substitutions on the phenyl ring selected from lower  
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are  
FT independently selected from hydrogen and lower alkyl"  
FT Modified-site 14  
FT /note= "Preferably C-terminus is -C(O)R2 where R2 is  
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3  
FT and R4 are independently selected from hydrogen and lower  
FT alkyl, and where the nitrogen atom of the -NR3R4 group  
FT can optionally be the amine group of the N-terminus of  
FT the peptide forming a cyclic peptide"  
XX  
PN WO640189-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 05-JUN-1996; 96WO-US008998.  
XX  
PR 07-JUN-1995; 95US-00472371.  
PR 07-JUN-1995; 95US-00473604.  
PR 07-JUN-1995; 95US-00476168.  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00484090.  
PR 07-JUN-1995; 95US-00485301.  
XX  
PA (GLAXO) GLAXO GROUP LTD.  
PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX  
DR WPI; 1997-051883/05.  
XX  
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide  
PT mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.  
XX  
PS Claim 18; Page 89; 106pp; English.  
XX  
CC The present sequence is a compound which binds to thrombopoietin (TPO)  
CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding  
CC affinity to TR as expressed by an IC50 of no more than about 100 nM. The  
CC compound (especially if modified, see features table) can be used for  
CC treating patients suffering from haematological disorders and  
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
CC marrow transfusions. The peptide may also be used to maintain the  
CC proliferation and growth of TPO-dependent cell lines and for use in  
CC biological research, for detecting TPO receptors on living cells  
XX



## SQ Sequence 14 AA:

Query Match 100.0%; Score 85; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 |||||  
 1 CADGPTLRWISFC 14

## RESULT 3

AAW09465  
 ID AAW09465 standard; protein; 14 AA.

AC AAW09465;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound cyclic peptide.

KM Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;

KM bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

FX Key Location/Qualifiers

FT Disulfide-bond 1..14

PN WO9640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

PS Claim 30; Page 91; 106pp; English.

CC The present sequence is a compound which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering

CC from haematological disorders and thrombocytopenia resulting from

CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide

CC may also be used to maintain the proliferation and growth of TPO-

CC dependent cell lines and for use in biological research, for detecting

CC TPO receptors on living cells

XX Sequence 14 AA;

Query Match 100.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 |||||  
 1 CADGPTLRWISFC 14

## RESULT 4

AAW09482  
 ID AAW09482 standard; protein; 14 AA.

AC AAW09482;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding peptide.

KM Haematology; thrombocytopenia; TPO; TR; proliferation;

KM bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

PN WO9640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 26; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering

CC from haematological disorders and thrombocytopenia resulting from

CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide

CC may also be used to maintain the proliferation and growth of TPO-

CC dependent cell lines and for use in biological research, for detecting

CC TPO receptors on living cells

XX Sequence 14 AA;

Query Match 100.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 |||||  
 1 CADGPTLRWISFC 14

## RESULT 5

AAW33031  
 ID AAW33031 standard; peptide; 14 AA.

AC AAW33031;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KM Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KW	radiation therapy; bone marrow transfusion; diagnosis;
KM	signal transduction; receptor activation; cell culture.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Disulfide-bond 1..14
XX	
PM	WO9640750-A1.
XX	
PD	19-DEC-1996.
XX	
PF	07-JUN-1996; 96WO-US009623.
XX	
PR	07-JUN-1995; 95US--00478128.
PR	07-JUN-1995; 95US-00485301.
XX	
PA	(GLAXO) GLAXO GROUP LTD.
PI	Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI	Matheakie LC, Schatz PJ, Magstrom CR, Wrighton NC;
XX	WPI, 1997-052226/05.
DR	
XX	
PT	Peptides and peptide mimetics which bind to and activate the
PT	thrombopoietin receptor - useful in treatment of haematological
XX	disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX	
PS	Claim 30; Page 91; 106pp; English.
XX	
CC	The present peptide binds the thrombopoietin receptor (TR), has a
CC	molecular weight of less than 8000 Da and a TR binding affinity as
CC	expressed by an IC50 of no more than about 100 microm.
CC	treat disorders which are susceptible to treatment with a thrombopoietin
CC	agonist, preferably haematological disorders and thrombocytopaenia
CC	resulting from chemotherapy, radiation therapy or bone marrow
CC	transfusions. It can also be used diagnostically, e.g. to investigate the
CC	mechanism of thrombopoietin signal transduction and receptor activation,
CC	or to maintain the proliferation and growth of thrombopoietin dependent
CC	cell lines
XX	
SO	Sequence 14 AA:
XX	
QY	Query Match 100.0%; Score 85; DB 2; Length 14;
XX	Best Local Similarity 100.0%; Pred. No. 6.9e-07;
DB	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX	
QY	1 CADGPTLRWISFC 14
XX	
DB	1 CADGPTLRWISFC 14
XX	
RESULT 6	
AAM36633	
ID	AAM36633 standard; peptide; 14 AA.
XX	
AC	AAM36633;
XX	
DT	11-MAR-1998 (first entry)
XX	
DE	Thrombopoietin receptor binding peptide.
XX	
KW	Thrombopoietin receptor; binding peptide; treatment; agonist;
KW	haematological disorder; thrombocytopaenia; chemotherapy;
KM	radiation therapy; bone marrow transfusion; diagnosis;
KM	signal transduction; receptor activation; cell culture.
XX	
OS	Synthetic.
XX	
PN	WO9640750-A1.
PD	19-DEC-1996.
XX	

Query Match	Best Local Similarity	100.0%; Score 85; DB 2; Length 14;
Matches 14; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY 1 CADGPTLRWISFC 14		
DB 1 CADGPTLRWISFC 14		
RESULT 7		
AAW33029		
ID AAW33029 standard; peptide; 14 AA.		
XX AAW33029;		
AC 11-MAR-1998 (first entry)		
DT Thrombopoietin receptor binding peptide.		
XX Thrombopoietin receptor; binding peptide; treatment; agonist;		
DE haematological disorder; thrombocytopaenia; chemotherapy;		
KW radiation therapy; bone marrow transplantation; diagnosis;		
KM signal transduction; receptor activation; cell culture.		
XX OS Synthetic.		
XX PN WO9640750-A1.		
PN 19-DEC-1996.		
XX PF 07-JUN-1996; 96WO-US009623.		
XX PR 07-JUN-1995; 95US-00478128.		
XX PR 07-JUN-1995; 95US-00485301.		
PA (GLAXO) GLAXO GROUP LTD.		
XX Dower WJ, Barrat RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;		
PI Mattheakis LC, Schatz PJ, Wagsstrom CR, Wrighton NC;		
XX WPI; 1997-052226/05.		
XX Peptides and peptide mimetics which bind to and activate the		
PT thrombopoietin receptor - useful in treatment of haematological		

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 19; Page 89; 106pp; English.

XX

CC The present peptide binds the thrombopoietin receptor (TR), has a

CC molecular weight of less than 8000 Da and a TR binding affinity as

CC expressed by an IC50 of no more than about 100 microm. It can be used to

CC treat disorders which are susceptible to treatment with a thrombopoietin

CC agonist, preferably haematological disorders and thrombocytopenia

CC resulting from chemotherapy, radiation therapy or bone marrow

CC transfusions. It can also be used diagnostically, e.g. to investigate the

CC mechanism of thrombopoietin signal transduction and receptor activation,

CC or to maintain the proliferation and growth of thrombopoietin dependent

CC cell lines

CC

SQ Sequence 14 AA;

Query Match 100.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14

Db 1 CADGPTLRWISFC 14

RESULT 8

AAW35401

ID AAW35401 standard; peptide; 14 AA.

XX

AC AAW35401;

XX

DT 11-MAR-1998 (first entry)

XX

DE Thrombopoietin receptor binding peptide.

XX

KM Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

KM signal transduction; receptor activation; cell culture.

XX

OS Synthetic.

XX

FX Key Location/Qualifiers

FT Disulfide-bond 1. 14

FT Modified-site 14 /note= "NH2-D-Cys"

XX

XX W09640750-A1.

XX

XX 19-DEC-1996.

XX

PF 07-JUN-1996; 96WO-US009623.

XX

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX

PA (GLAXO) GLAXO GROUP LTD.

XX

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX

XX WPI; 1997-052226/05.

XX

PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX

XX Example 6; Page 63; 106pp; English.

XX

CC The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14

Db 1 CADGPTLRWISFC 14

RESULT 9

AAW36647

ID AAW36647 standard; peptide; 14 AA.

XX

AC AAW36647;

XX

DT 11-MAR-1998 (first entry)

XX

DE Thrombopoietin receptor binding peptide.

XX

KM Thrombopoietin receptor; binding peptide; treatment; agonist;

KM haematological disorder; thrombocytopenia; chemotherapy;

KM radiation therapy; bone marrow transfusion; diagnosis;

KM signal transduction; receptor activation; cell culture.

XX

OS Synthetic.

XX

FX Key Location/Qualifiers

FT Disulfide-bond 1. 14

FT Modified-site 14 /note= "NH2-D-Cys"

XX

XX W09640750-A1.

XX

XX 19-DEC-1996.

XX

PF 07-JUN-1996; 96WO-US009623.

XX

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX

PA (GLAXO) GLAXO GROUP LTD.

XX

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX

XX WPI; 1997-052226/05.

XX

PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX

XX Disclosure; Page 26; 106pp; English.

XX

CC The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 85; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14

```

Db          1 CADGPTLRWISFC 14

RESULT 10
AAW35400 standard; peptide; 14 AA.
XX
XX AAW35400;
XX
XX 11-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopaenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1. .14
XX Misc-difference 1 /note= "D-form residue"
XX Modified-site 14 /note= "NH2-D-Cys"
XX
XX MO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
XX Example 6; Page 63; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 85; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 6.9e-07;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CADGPTLRWISFC 14
XX 1 CADGPTLRWISFC 14
XX
XX RESULT 11
XX AAW33032
XX ID AAW33032 standard; peptide; 14 AA.
XX

```

```

AC AAW33032;
XX
XX 11-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopaenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1. .14
XX Modified-site 1 /note= "acylated"
XX Modified-site 14 /note= "amidated"
XX
XX MO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
XX Claim 30; Page 91; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
XX molecular weight of less than 8000 Da and a TR binding affinity as
XX expressed by an IC50 of no more than about 100 microm. It can be used to
XX treat disorders which are susceptible to treatment with a thrombopoietin
XX agonist, preferably haematological disorders and thrombocytopaenia
XX resulting from chemotherapy, radiation therapy or bone marrow
XX transfusions. It can also be used diagnostically, e.g. to investigate the
XX mechanism of thrombopoietin signal transduction and receptor activation,
XX or to maintain the proliferation and growth of thrombopoietin dependent
XX cell lines
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 85; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 6.9e-07;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CADGPTLRWISFC 14
XX 1 CADGPTLRWISFC 14
XX
XX AAB17014
XX ID AAB17014 standard; peptide; 14 AA.
XX
XX AAB17014;
XX
XX 31-OCT-2000 (first entry)
XX
XX TPO-mimetic peptide sequence SEQ ID NO:70.
XX

```

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KM autoimmunity disease; cytostatic; antineoplastic; thrombolytic; VEGF;  
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KM inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KM vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KM thrombosis; pharmaceutical.  
 OS Synthetic.  
 OS  
 PN WO200024782-A2.  
 XX  
 PD 04-MAY-2000.  
 XX  
 PF 25-OCT-1999; 99WO-US025044.  
 XX  
 PR 23-OCT-1998; 98US-0105371P.  
 XX  
 PR 22-OCT-1999; 99US-00428082.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Feige U, Liu C, Cheetham J, Boone TC;  
 XX  
 DR WPI; 2000-350702/30.  
 XX  
 PT Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 PS Claim 19; Page 218; 608pp; English.  
 XX  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from: -(L1)-C-P1, -(L1)-C-P1-(L2)d-P2, -(L1)-C-P1-  
 CC (L2)d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antineoplastic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AA69443 to AA69526 and AA61955 to  
 CC AA61803 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 100.0%; Score 85; DB 3; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 Db 1 CADGPTLRWISFC 14  
 XX  
 RESULT 13  
 ID AAV25826 standard; peptide; 14 AA.  
 XX  
 AC AAV25826;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #12.  
 XX  
 KM Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KM hemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

KM bone marrow transplantation; haematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
 OS Homo sapiens.  
 OS  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 XX  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PR 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwila SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagsstrom CR, Hendren RW, Deprience RB, Poddaturi S;  
 PI Yarn Q;  
 XX  
 DR WPI; 2001-564142/53.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 67-68; 128pp; English.  
 XX  
 CC Sequences AAV25815-AAV26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent hematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 100.0%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 Db 1 CADGPTLRWISFC 14  
 XX  
 RESULT 14  
 ID AAV25852 standard; peptide; 14 AA.  
 XX  
 AC AAV25852;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #38.  
 XX  
 KM Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Depirince RB, Podduturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 DB 1 CADGPTLRWISFC 14  
 RESULT 15  
 AAU25866  
 ID AAU25866 standard; peptide; 14 AA.  
 XX  
 AC AAU25866;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human thrombopoietin receptor (TPO-R) activator peptide #52.  
 XX

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6251864-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 01-MAR-2000; 2000US-00516704.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Depirince RB, Podduturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 PS Disclosure; Col 20; 128pp; English.  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 85; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 DB 1 CADGPTLRWISFC 14  
 RESULT 16  
 ABB72900  
 ID ABB72900 standard; peptide; 14 AA.  
 XX  
 AC ABB72900;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:70.  
 XX

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor; alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antirheumatic; antiarthritic; haemostatic; dermatological;  
 KW antihaematic; anorectic; antifertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 KM  
 XX Homo sapiens.  
 OS Synthetic.  
 OS  
 XX WO200183525-A2.  
 PN  
 XX 08-NOV-2001.  
 PD  
 XX 02-MAY-2001; 2001WO-US014310.  
 PF  
 XX 03-MAY-2000; 2000US-00563286.  
 PR  
 PA (AMGE-) AMGEN INC.  
 PI  
 XX Felge U, Liu C, Cheetham JC, Boone TC, Gudas JM;  
 PI WPI; 2002-130313/17.  
 DR  
 XX Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 PS  
 XX Claim 39; Page 44; 176pp; English.  
 PS  
 XX The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antinaeumatic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL25695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 85; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CADGPTLRWISFC 14  
 |||||  
 Db 1 CADGPTLRWISFC 14

ID ADJ73051 standard; peptide; 14 AA.  
 XX  
 AC ADJ73051;  
 XX  
 XX 06-MAY-2004 (first entry)  
 DT  
 XX TPO mimetic peptide sequence SeqID 505.  
 DE  
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;  
 KW TPO.  
 KM  
 XX  
 XX Synthetic.  
 OS  
 XX WO2003084477-A2.  
 PN  
 XX 16-OCT-2003.  
 PD  
 XX 24-MAR-2003; 2003WO-US009139.  
 PF  
 XX 29-MAR-2002; 2002US-0368791P.  
 PR  
 XX (CENZ ) CENTOCOR INC.  
 PA  
 XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;  
 PI WPI; 2003-804237/75.  
 DR  
 XX New CDR mimetibody comprising a portion of a heavy or light chain  
 PT variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PT cardiovascular or neurologic disease.  
 PS  
 XX Disclosure; SEQ ID NO 505; 97pp; English.  
 PS  
 XX This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LBR). The present invention describes human  
 CC mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neurologic  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 CC  
 XX  
 SQ Sequence 14 AA;  
 Query Match 100.0%; Score 85; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CADGPTLRWISFC 14  
 |||||  
 Db 1 CADGPTLRWISFC 14

RESULT 18  
 ADJ52686  
 ID ADJ52686 standard; peptide; 14 AA.  
 XX  
 AC ADJ52686;  
 XX  
 XX 06-MAY-2004 (first entry)  
 DT  
 XX CHI deleted mimetibody-related peptide SeqID505.  
 DE

XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiac;  
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;  
 KW arrhythmia; hypertension; heart failure; neurodegenerative;  
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
 KW cancerous condition; infectious disease; bacterial infection;  
 KW viral infection; fungal infection.  
 XX Unidentified.  
 OS Synthetic.  
 PN WO2004002417-A2.  
 XX PD 08-JAN-2004.  
 XX PF 27-JUN-2003; 2003WO-US020347.  
 XX PR 28-JUN-2002; 2002US-0392431P.  
 XX (GEN2 ) CENTOCOR INC.  
 PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nespore TC;  
 PI Kutolowski KA;  
 DR WPI; 2004-082870/08.  
 XX PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for  
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
 PT diseases.  
 XX PS Claim 2; SEQ ID NO 505; 129pp; English.  
 XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an immunosuppressive,  
 CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,  
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody  
 CC is useful for diagnosing or treating a disease condition in a cell,  
 CC tissue, organ or animal, specifically for modulating, treating,  
 CC alleviating, preventing the incidence or reducing the symptoms of an  
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
 CC conditions, or infectious diseases (for example bacterial, viral or  
 CC fungal infection). The present sequence is that of a peptide which may be  
 CC used during the creation of a mimetibody of the invention.  
 XX SQ Sequence 14 AA;  
 XX  
 XX Query Match 100.0%; Score 85; DB 8; Length 14;  
 XX Best Local Similarity 100.0%; Pred. No. 6, 9e-07;  
 XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 Db 1 CADGPTLRWISFC 14  
 RESULT 19  
 ADJ51647  
 ID ADJ51647 standard; peptide; 14 AA.  
 XX  
 XX AC ADJ51647;  
 XX DT 06-MAY-2004 (first entry)  
 XX CH1 deleted mimetibody-related peptide SeqID505.  
 XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
 KW anti-allergic; muscular-Gen; cytosstatic; anti-inflammatory; neuroleptic;  
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;  
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
 KW obstetric disorder; haematological disorder; immunological disorder;  
 KW allergic disorder; infectious disorder; musculoskeletal disorder;  
 KW oncological disorder; neurological disorder; nutritional disorder;  
 KW ophthalmological disorder; pediatric disorder; psychiatric disorder;  
 KW renal disorder; pulmonary disorder.  
 XX Unidentified.  
 OS Synthetic.  
 PN WO2004002424-A2.  
 XX PD 08-JAN-2004.  
 XX PF 30-JUN-2003; 2003WO-US020495.  
 XX PR 28-JUN-2002; 2002US-0392431P.  
 XX PR 19-SEP-2002; 2002US-0412144P.  
 XX (GEN2 ) CENTOCOR INC.  
 PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nespore TC;  
 PI Kutolowski KA;  
 DR WPI; 2004-082872/08.  
 XX PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for  
 PT diagnosing, preventing or treating cardiovascular, dermatologic and  
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
 PT nutritional disorders.  
 XX PS Claim 15; SEQ ID NO 505; 123pp; English.  
 XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences  
 CC which encode them), compositions, methods and uses. The invention may be  
 CC useful for the development of compounds with an osteopathic,  
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
 CC immunomodulator, anti-allergic, ophthalmological, nephrotropic or  
 CC anti-inflammatory, neuroleptic, acting as a tumour necrosis factor (TNF)-  
 CC modulator or cytokine-agonist. The methods and compositions of the  
 CC present invention are useful for the diagnosis, prevention and/or  
 CC treatment of diseases or conditions associated with aberrant expression  
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,  
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
 CC obstetric, haematologic, immunological, allergic, infectious,  
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
 CC pediatric, psychiatric, renal or pulmonary disorders. The present  
 CC sequence is that of a peptide which may be used during the creation of a  
 CC mimetibody of the invention.  
 XX SQ Sequence 14 AA;  
 XX  
 XX Query Match 100.0%; Score 85; DB 8; Length 14;  
 XX Best Local Similarity 100.0%; Pred. No. 6, 9e-07;  
 XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 Db 1 CADGPTLRWISFC 14  
 RESULT 20  
 AAM09456  
 ID AAM09456 standard; protein; 18 AA.



XX	AAM09456;
AC	
XX	10-SEP-1997 (first entry)
DT	
XX	Thrombopoietin receptor binding compound peptide.
DE	
XX	Haematology; thrombocytopenia; TPO; TR; proliferation; bone marrow transfusion; chemotherapy; radiation therapy.
KM	
XX	Synthetic.
OS	
XX	
Key	Location/Qualifiers
FH	1, 18
FT	/note= "Preferably linkages are selected from: - CH2OC(O)NR;- phosphonate; -CH2S(O)ZNR;- -C(O)NR6 ;-NHC(O)NH; where R is hydrogen or lower alkyl and R6 is lower alkyl"
FT	1
Modified-site	/note= "Preferably N-terminus is selected from: -NRRL; -NRC(O)R;-NRC(O)OR;-NRS(O)Z;-NHC(O)NHR; succinimide; benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3 substitutions on the phenyl ring selected from lower alkyl, lower alkoxy, chloro, bromo; where R and R1 are independently selected from hydrogen and lower alkyl"
FT	18
Modified-site	/note= "Preferably C-terminus is -C(O)R2 where R2 is selected from hydroxy, lower alkoxy, and -NR3R4, where R3 and R4 are independently selected from hydrogen and lower alkyl, and where the nitrogen atom of the -NR3R4 group can optionally be the amine group of the N-terminus of the peptide forming a cyclic peptide"
FT	
PN	WO9640189-A1.
PD	19-DEC-1996.
PP	05-JUN-1996; 96WO-US008998.
PR	07-JUN-1995; 95US-00472371.
PR	07-JUN-1995; 95US-00473604.
PR	07-JUN-1995; 95US-00476168.
PR	07-JUN-1995; 95US-00478128.
PR	07-JUN-1995; 95US-00484090.
PR	07-JUN-1995; 95US-00485301.
PA	(GLAXO ) GLAXO GROUP LTD.
P1	Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS; Matheakis LC, Schatz PJ, Wagelstrom CR, Wrighton NC; WPI, 1997-051883/05.
PT	Thrombopoietin receptor-binding/activating peptide(s) and peptide mimetic(s) - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
PS	Claim 18; Page 89; 106pp; English.
XX	The present sequence is a compound which binds to thrombopoietin (TPO) receptor (TR). It has a molecular weight of < 8000 Da, and a binding affinity to TR as expressed by an IC50 of no more than about 100 nM. The compound (especially if modified, see features table) can be used for treating patients suffering from haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transplants. The peptide may also be used to maintain the proliferation and growth of TPO-dependent cell lines and for use in biological research, for detecting TPO receptors on living cells
XX	Sequence 18 AA:
Query Match	100.0%; Score 85; DB 2; Length 18;
Best Local Similarity	100.0%; Ptd. No. 9e-07;

Matches	14; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	1 CADGPTLRWISFC 14							
Db	3 CADGPTLRWISFC 16							
RESULT 21								
ID	AAW33023							
	AAW33023 standard; peptide; 18 AA.							
AC	AAW33023;							
XX								
DT	11-MAR-1998 (first entry)							
DE	Thrombopoietin receptor binding peptide.							
XX								
KW	Thrombopoietin receptor; binding peptide; treatment; agonist;							
KM	haematological disorder; thrombocytopaenia; chemotherapy;							
KW	radiation therapy; bone marrow transfusion; diagnosis;							
KW	signal transduction; receptor activation; cell culture.							
OS	Synthetic.							
XX								
PN	WO9640750-A1.							
PD	19-DEC-1996.							
XX								
PF	07-JUN-1996; 96WO-US009623.							
XX								
PR	07-JUN-1995; 95US-00478128.							
PR	07-JUN-1995; 95US-00485301.							
XX								
PA	(GLAXO) GLAXO GROUP LTD.							
PI	Dower MJ, Barret RM, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;							
PI	Matheakis LC, Schatz PJ, Wagerstrom CR, Wrighton NC;							
XX								
DR	WPI; 1997-052226/05.							
XX								
PT	Peptides and peptide mimetics which bind to and activate the							
PT	thrombopoietin receptor - useful in treatment of haematological							
PT	disorders, esp. thrombocytopenia resulting from chemotherapy, etc.							
XX								
PS	Claim 19; Page 89; 106pp; English.							
XX								
CC	The present peptide binds the thrombopoietin receptor (TR), has a							
CC	molecular weight of less than 8000 Da and a TR binding affinity as							
CC	expressed by an IC50 of no more than about 100 microm, it can be used to							
CC	treat disorders which are susceptible to treatment with a thrombopoietin							
CC	agonist, preferably haematological disorders and thrombocytopaenia							
CC	resulting from chemotherapy, radiation therapy or bone marrow							
CC	transfusions. It can also be used diagnostically, e.g. to investigate the							
CC	mechanism of thrombopoietin signal transduction and receptor activation,							
CC	or to maintain the proliferation and growth of thrombopoietin dependent							
CC	cell lines							
XX								
SQ	Sequence 18 AA;							
Oy	Query Match	100.0%;	Score 85;	DB 2;	Length 18;			
	Best Local Similarity	100.0%;	Pred. NO. 9e-07;					
	Matches 14; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	1 CADGPTLRWISFC 14							
Db	3 CADGPTLRWISFC 16							
RESULT 22								
ID	AAAB17020							

XX 31-OCT-2000 (first entry)  
 DT TPO-mimetic peptide sequence SEQ ID NO:76.  
 XX  
 DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTAA4; mimetic; IL-1; TNF; antagonist; MMP;  
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
 KW thrombosis; pharmaceutical.  
 XX Synthetic.  
 OS  
 XX WO200024782-A2.  
 PN  
 XX 04-MAY-2000.  
 PD  
 XX 25-OCT-1999; 99WO-US025044.  
 PF  
 XX 23-OCT-1998; 98US-0105371P.  
 PR 22-OCT-1999; 98US-00428082.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Feige U, Liu C, Cheetham J, Boone TC;  
 XX  
 DR WPI; 2000-350702/30.  
 XX  
 PT Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides; useful for treating cancer and autoimmune diseases.  
 XX  
 XX Claim 19; Page 220; 608pp; English.  
 PS  
 XX  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-F1-(X2)-b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)-d-P2, -(L1)-c-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AA69443 to AA69526 and AA61955 to  
 CC AA61803 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 18 AA;  
 Query Match 100.0%; Score 85; DB 3; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 DB 3 CADGPTLRWISFC 16  
 AAU25820  
 ID AAU25820 standard; peptide; 18 AA.  
 XX  
 AC AAU25820;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX

DE Human thrombopoietin receptor (TPO-R) activator peptide #6.  
 XX  
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US6251864-B1.  
 PN  
 XX  
 PD 26-JUN-2001.  
 PF  
 XX 01-MAR-2000; 2000US-00516704.  
 PR  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 PR 07-JUN-1996; 96WO-US009623.  
 PR 15-AUG-1996; 96US-00699027.  
 XX  
 PA (GLAXO) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ,  
 PI Balasubramanian P, Wagstroom CR, Hendren RW, Deprince RB, Podduturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 2001-564142/63.  
 XX  
 PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.  
 XX  
 XX Disclosure; Col 65-66; 128pp; English.  
 PS  
 XX  
 CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines  
 CC  
 XX  
 SQ Sequence 18 AA;  
 Query Match 100.0%; Score 85; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CADGPTLRWISFC 14  
 DB 3 CADGPTLRWISFC 16  
 ABB72906  
 ID ABB72906 standard; peptide; 18 AA.  
 XX  
 AC ABB72906;  
 XX  
 DT 05-APR-2002 (first entry)  
 XX

XX TPO mimetic peptide SEQ ID NO:76.  
 DE  
 XX  
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;  
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KM TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KM TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KM MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antineoplastic; antidiabetic; opthalmological;  
 KM antianemic; anorectic; antifertility; haemostatic; dermatological;  
 KM neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KM cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KM sleep disorder; neurological degenerative disease; anaemia;  
 KM thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KM Fanconi's syndrome.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 XX WO200183525-A2.  
 PN  
 XX 08-NOV-2001.  
 PD  
 XX 02-MAY-2001; 2001WO-US014310.  
 PF  
 XX 03-MAY-2000; 2000US-00563286.  
 PR  
 XX (AMGE-) AMGEN INC.  
 PA  
 XX  
 PI Feige U, Lhu C, Cheetham JC, Boone TC, Gudas JM;  
 XX  
 DR WPI; 2002-130313/17.  
 XX  
 PT Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility.  
 XX  
 PS Claim 39; Page 44; 176pp; English.  
 XX  
 XX The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimer, (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cyrostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL5695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 XX Sequence 18 AA;  
 SQ  
 Query Match 100.0%; Score 85; DB 5; Length 18;  
 Best Local Similarity 100.0%; Pred. NO. 9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 25  
 ADJ73058  
 ID ADJ73058 standard; peptide; 18 AA.  
 XX  
 XX  
 AC ADJ73058;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 XX TPO mimetic peptide sequence SeqID 512.  
 DE  
 XX  
 XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;  
 KM cardiovascular; infectious; malignant; neurologic disease; anaemia;  
 KM immunomodulator; cardiant; antimicrobial; cyostatic; neuroprotective;  
 KM TPO.  
 KM  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX WO2003084477-A2.  
 PN  
 XX 16-OCT-2003.  
 PD  
 XX 24-MAR-2003; 2003WO-US009139.  
 PF  
 XX 29-MAR-2002; 2002US-0368791P.  
 PR  
 XX (CBNZ ) CENTOCOR INC.  
 PA  
 XX  
 PI Heavner GA, Knight DW, Scallion BJ, Ghayeb J;  
 XX  
 DR WPI; 2003-804237/75.  
 XX  
 PT New CDR mimetibody comprising a portion of a heavy or light chain  
 PT variable region comprising human framework or ligand binding region,  
 PT useful for preparing a composition for treating e.g., immune,  
 PT cardiovascular or neurologic disease.  
 XX  
 PS Disclosure; SEQ ID NO 512; 97pp; English.  
 XX  
 XX This invention relates to novel mammalian CDR mimetibodies, specific  
 CC portions or variants thereof. Specifically, it refers to an antibody  
 CC fragment where a protein has been inserted into, or replaces a portion  
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at  
 CC least one portion of a heavy chain or light chain variable region, which  
 CC itself comprises at least one human framework region and at least one  
 CC ligand binding region (LBR). The present invention describes human  
 CC mimetibodies, including modified immunoglobulins and cleavage products  
 CC that can be useful in gene therapy and the generation of transgenic  
 CC plants and animals. Furthermore, the CDR mimetibody is useful for  
 CC preparing compositions for modulating, treating or reducing the symptoms  
 CC of immune, cardiovascular, infectious, malignant and/or neurologic  
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,  
 CC cardiant, antimicrobial, cyostatic and neuroprotective activities. This  
 CC peptide sequence is a TPO mimetic peptide sequence used to make a  
 CC mimetibody of the invention.  
 CC  
 XX  
 XX Sequence 18 AA;  
 SQ  
 Query Match 100.0%; Score 85; DB 7; Length 18;  
 Best Local Similarity 100.0%; Pred. NO. 9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 26  
 ADJ52693  
 ID ADJ52693 standard; peptide; 18 AA.  
 XX  
 XX  
 AC ADJ52693;  
 XX  
 DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID512.  
DE  
XX  
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiac;  
XX hypotensive; neuroprotective; nootropic; antibacterial; virucide;  
XX functional; gene therapy; immune disorder; cardiovascular disease;  
XX arrhythmia; hypertension; heart failure; neurodegenerative;  
XX multiple sclerosis; dementia; Alzheimer's disease; anaemia;  
XX cancerous condition; infectious disease; bacterial infection;  
XX viral infection; fungal infection.  
XX  
XX Unidentified.  
OS Synthetic.  
PN WO2004002417-A2.  
XX  
XX 08-JAN-2004.  
XX  
XX 27-JUN-2003; 2003WO-US020347.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX  
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;  
PI Kutoloski KA;  
XX  
XX WPI; 2004-082870/08.  
XX  
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for  
PT modulating, treating, alleviating, preventing an immune, cardiovascular,  
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious  
PT diseases.  
XX  
XX Claim 2; SEQ ID NO 512; 129pp; English.  
PS  
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an immunosuppressive,  
CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,  
CC antibacterial, virucide or fungicide activity. In addition, the disclosed  
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody  
CC is useful for diagnosing or treating a disease condition in a cell,  
CC tissue, organ or animal, specifically for modulating, treating, or  
CC alleviating, preventing the incidence or reducing the symptoms of an  
CC immune, cardiovascular (for example arrhythmia, hypertension or heart  
CC failure), or neurodegenerative (for example multiple sclerosis, dementia  
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous  
CC conditions, or infectious diseases (for example bacterial, viral or  
CC fungal infection). The present sequence is that of a peptide which may be  
CC used during the creation of a mimetibody of the invention.  
CC  
XX  
SQ Sequence 18 AA;  
Query Match 100.0%; Score 85; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. NO. 9e-07; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;  
OY 1 CADGPTLRWISFC 14  
|||  
DB 3 CADGPTLRWISFC 16  
|||  
RESULT 27  
ADJ51654  
ID ADJ51654 standard; peptide; 18 AA.  
XX  
XX ADJ51654;  
AC  
XX  
XX 06-MAY-2004 (first entry)  
DT  
XX  
XX CH1 deleted mimetibody-related peptide SeqID512.  
DE

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;  
XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;  
XX gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;  
XX anti-allergic; muscular-Gen; cytosstatic; anti-inflammatory; neuroleptic;  
XX ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;  
XX TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;  
XX dental disorder; oral disorder; dermatological disorder; ear disorder;  
XX nose disorder; throat disorder; endocrine disorder; metabolic disorder;  
XX gastrointestinal disorder; gynaecological disorder; hepatic disorder;  
XX obstetric disorder; haematologic disorder; immunological disorder;  
XX allergic disorder; infectious disorder; musculoskeletal disorder;  
XX oncological disorder; neurological disorder; nutritional disorder;  
XX ophthalmologic disorder; pediatric disorder; psychiatric disorder;  
XX renal disorder; pulmonary disorder.  
XX  
XX Unidentified.  
OS Synthetic.  
PN WO2004002424-A2.  
XX  
XX 08-JAN-2004.  
XX  
XX 30-JUN-2003; 2003WO-US020495.  
XX  
XX 28-JUN-2002; 2002US-0392431P.  
XX  
XX 19-SEP-2002; 2002US-0412144P.  
XX  
XX (CENZ ) CENTOCOR INC.  
XX  
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;  
PI Kutoloski KA;  
XX  
XX WPI; 2004-082872/08.  
XX  
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for  
PT diagnosing, preventing or treating cardiovascular, dermatologic,  
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and  
PT nutritional disorders.  
XX  
XX Claim 15; SEQ ID NO 512; 123pp; English.  
PS  
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences  
CC which encode them), compositions, methods and uses. The invention may be  
CC useful for the development of compounds with an osteopathic,  
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,  
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,  
CC immunomodulator, anti-allergic, muscular-Gen, cytosstatic,  
CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or  
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-  
CC modulator or cytokine-agonist. The methods and compositions of the  
CC present invention are useful for the diagnosis, prevention and/or  
CC treatment of diseases or conditions associated with aberrant expression  
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,  
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,  
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,  
CC obstetric, haematologic, immunological, allergic, infectious,  
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,  
CC pediatric, psychiatric, renal or pulmonary disorders. The present  
CC sequence is that of a peptide which may be used during the creation of a  
CC mimetibody of the invention.  
CC  
XX  
SQ Sequence 18 AA;  
Query Match 100.0%; Score 85; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. NO. 9e-07; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;  
OY 1 CADGPTLRWISFC 14  
|||  
DB 3 CADGPTLRWISFC 16  
|||  
RESULT 28

AAW09458  
ID AAW09458 standard; protein; 19 AA.  
XX  
AC AAW09458;  
XX  
DT 10-SEP-1997 (first entry)  
XX  
DE Thrombopoietin receptor binding compound peptide.  
XX  
KM Haematology; thrombocytopenia; TPO; TR; proliferation;  
XX bone marrow transfusion; chemotherapy; radiation therapy.  
XX  
OS Synthetic.  
XX  
FH Key  
FH MISC-difference 1. 19  
FT /note= "Preferably linkages are selected from: -  
FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6  
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is  
FT lower alkyl"  
FT Modified-site  
FT 1  
FT /note= "Preferably N-terminus is selected from: -NRR1; -  
FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NR; succinimide;  
FT benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3  
FT substitutions on the phenyl ring selected from lower  
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are  
FT independently selected from hydrogen and lower alkyl"  
FT 19  
FT /note= "Preferably C-terminus is -C(O)R2 where R2 is  
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3  
FT and R4 are independently selected from hydrogen and lower  
FT alkyl, and where the nitrogen atom of the -NR3R4 group  
FT can optionally be the amine group of the N-terminus of  
FT the peptide forming a cyclic peptide"  
XX  
PN W09640189-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 05-JUN-1996; 96WO-US008998.  
XX  
PR 07-JUN-1995; 95US-004722371.  
PR 07-JUN-1995; 95US-00473604.  
PR 07-JUN-1995; 95US-00476168.  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00484090.  
PR 07-JUN-1995; 95US-00485301.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX  
DR WPI; 1997-051883/05.  
XX  
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide  
PT mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.  
XX  
PS Claim 18; Page 89; 106pp; English.  
XX  
XX The present sequence is a compound which binds to thrombopoietin (TPO)  
XX receptor (TR). It has a molecular weight of < 8000 Da, and a binding  
XX affinity to TR as expressed by an IC50 of no more than about 100 mM. The  
XX compound (especially if modified, see features table) can be used for  
XX treating patients suffering from haematological disorders and  
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
XX marrow transfusions. The peptide may also be used to maintain the  
XX proliferation and growth of TPO-dependent cell lines and for use in  
XX biological research, for detecting TPO receptors on living cells  
XX  
SQ Sequence 19 AA;

Query Match 100.0%; Score 85; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Dy 1 CADGPTLRWISFC 14  
Db 3 CADGPTLRWISFC 16  
RESULT 29  
AAW33025  
ID AAW33025 standard; peptide; 19 AA.  
XX  
AC AAW33025;  
XX  
DT 11-MAR-1998 (first entry)  
XX  
DE Thrombopoietin receptor binding peptide.  
XX  
KM Thrombopoietin receptor; binding peptide; treatment; agonist;  
KM haematological disorder; thrombocytopenia; chemotherapy;  
KM radiation therapy; bone marrow transfusion; diagnosis;  
KM signal transduction; receptor activation; cell culture.  
XX  
OS Synthetic.  
XX  
PN W09640750-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 07-JUN-1996; 96WO-US009623.  
XX  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
XX  
DR WPI; 1997-052226/05.  
XX  
PT Peptides and peptide mimetics which bind to and activate the  
PT Thrombopoietin receptor - useful in treatment of haematological  
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
XX  
PS Claim 19; Page 89; 106pp; English.  
XX  
XX The present peptide binds the thrombopoietin receptor (TR), has a  
XX molecular weight of less than 8000 Da and a TR binding affinity as  
XX expressed by an IC50 of no more than about 100 microm. It can be used to  
XX treat disorders which are susceptible to treatment with a thrombopoietin  
XX agonist, preferably haematological disorders and thrombocytopenia  
XX resulting from chemotherapy, radiation therapy or bone marrow  
XX transfusions. It can also be used diagnostically, e.g. to investigate the  
XX mechanism of thrombopoietin signal transduction and receptor activation,  
XX or to maintain the proliferation and growth of thrombopoietin dependent  
XX cell lines  
XX  
SQ Sequence 19 AA;  
XX  
XX Query Match 100.0%; Score 85; DB 2; Length 19;  
XX Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Dy 1 CADGPTLRWISFC 14  
Db 3 CADGPTLRWISFC 16  
RESULT 30  
AAU25822  
ID AAU25822 standard; peptide; 19 AA.

```

XX AC AAU25822;
XX DT 17-DEC-2001 (first entry)
XX DE Human thrombopoietin receptor (TPO-R) activator peptide #8.
XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
XX KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
XX KW bone marrow transplantation; haematological disorder; platelet disorder;
XX KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
XX KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
XX KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX OS Homo sapiens.
XX PN US6251864-B1.
XX PD 26-JUN-2001.
XX PF 01-MAR-2000; 2000US-00516704.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PR 07-JUN-1996; 96WO-US000623.
XX PR 15-AUG-1996; 96US-00699027.
XX PA (GLAXO ) GLAXO GROUP LTD.
XX PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ,
XX PI Balasubramanian P, Magstrom CR, Hendren RM, Deprince RB, Podduturi S,
XX PI Yin Q,
XX DR WPI; 2001-564142/63.
XX PT Activating thrombopoietin receptors in cells, used to treat
XX PT thrombocytopenia and hematological disorders, comprises contacting cells
XX PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX PS Disclosure; Col 67-68; 128pp; English.
XX XX
XX XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX CC of activating thrombopoietin receptors in cells comprise contacting the
XX CC cells with effective amounts of peptides and peptide mimetics attached to
XX CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX CC as that due to chemotherapy, radiation therapy or bone-marrow
XX CC transplantation and to prevent thrombocytopenia in patients at risk. The
XX CC sequences are used to treat and prevent haematological disorders
XX CC including thrombocytopenia and platelet disorders. They are used in vitro
XX CC as unique tools for understanding the biological role of thrombopoietin
XX CC (TPO) and to develop other compounds that bind to and activate the TPO
XX CC receptor. The peptides can be used to detect TPO receptors on living
XX CC cells and fixed cells, in biological fluids, in tissue homogenates, and
XX CC in purified or natural biological materials. They may also be used for in
XX CC situ staining, fluorescence-activated cell sorting, Western blotting and
XX CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX CC be used for in vitro expansion of megakaryocytes and their committed
XX CC progenitors alone or in conjunction with additional cytokines
XX SQ Sequence 19 AA;
XX
XX Query Match 100.0%; Score 85; DB 4; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 9.6e-07;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

ID AAW09467 standard; protein; 13 AA.
XX AC AAW09467;
XX DT 10-SEP-1997 (first entry)
XX DE Thrombopoietin receptor binding compound cyclic peptide.
XX KW Haematology; thrombocytopenia; TPO; TR; proliferation; cyclic;
XX KW bone marrow transfusion; chemotherapy; radiation therapy.
XX OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT /note= "The Ala is linked with the modified Cys at
XX FT position 13"
XX FT Modified-site 14
XX FT /label= OTHER
XX FT /note= "S-carboxymethyl-cysteine alpha-carboxamide; the
XX FT forming a linkage onto the Ala at position one with the
XX FT delta C of this residue"
XX PN WO9640189-A1.
XX PD 19-DEC-1996.
XX PF 05-JUN-1996; 96WO-US0008998.
XX PR 07-JUN-1995; 95US-00472371.
XX PR 07-JUN-1995; 95US-00473604.
XX PR 07-JUN-1995; 95US-00476168.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00484090.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAXO ) GLAXO GROUP LTD.
XX PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Magstrom CR, Wighthon NC;
XX DR WPI; 1997-051883/05.
XX PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX PT mimetic(s) - useful in treatment of haematological disorders, esp.
XX PT thrombocytopenia resulting from chemotherapy, etc.
XX PS Claim 30, Page 91; 106pp; English.
XX XX
XX XX The present sequence is a compound which binds to thrombopoietin (TPO)
XX CC receptor (TR). The compound can be used for treating patients suffering
XX CC from haematological disorders and thrombocytopenia resulting from
XX CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
XX CC may also be used to maintain the proliferation and growth of TPO-
XX CC dependent cell lines and for use in biological research, for detecting
XX CC TPO receptors on living cells
XX XX
XX SQ Sequence 13 AA;
XX
XX Query Match 89.4%; Score 76; DB 2; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-05;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Db 2 AADGPTLRWISFC 14
XX 1 AADGPTLRWISFC 13
XX
XX RESULT 32
XX ID AAW35399 standard; peptide; 13 AA.
XX AC AAW35399;

```

```

XX 11-MAR-1998 (first entry)
DT Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
DE haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transplantation; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "COCH2-alanine linked via CH2 group to Cys13"
FT Modified-site 13
FT /note= "NH2-cytosine linked via sulfoxidised thiol group
FT to Ala1"
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96MO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Example 6; Page 63; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX
XX Sequence 13 AA;
SQ
Query Match 89.4%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ADGPTLRWISFC 14
DB 1 ADGPTLRWISFC 13

```

```

KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Cross-links 1
FT /note= "linked via disulfide bond to Cys1 of identical
FT peptide"
FT Modified-site 13
FT /note= "NH2-Phe"
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96MO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Example 9; Page 73; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX
XX Sequence 13 AA;
SQ
Query Match 89.4%; Score 76; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CADGPTLRWISF 13
DB 1 CADGPTLRWISF 13

```

```

RESULT 33
AAW35417
ID AAW35417 standard; peptide; 13 AA.
XX
XX AAW35417;
AC
XX
XX 11-MAR-1998 (first entry)
DT
XX
XX Thrombopoietin receptor binding peptide.
DE
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;

```

```

RESULT 34
AAW33033
ID AAW33033 standard; peptide; 13 AA.
XX
XX AAW33033;
AC
XX
XX 11-MAR-1998 (first entry)
DT
XX
XX Thrombopoietin receptor binding peptide.
DE
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1

```

[illegible]



XX Dower WJ, Barret RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 6; Page 64; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 XX  
 SQ Sequence 13 AA;  
 SO  
 Query Match 89.4%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ADGPTLRWISFC 14  
 Db 1 ADGPTLRWISFC 13  
 RESULT 37  
 ID AAW35422 standard; peptide; 13 AA.  
 XX  
 AC AAW35422;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "optionally acylated"  
 FT Cross-links 13 /note= "linked via disulfide bond to Cys13 of identical  
 FT peptide"  
 XX  
 FT W09640750-A1.  
 XX  
 PN 19-DEC-1996.  
 PD  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 XX Example 9; Page 74; 106pp; English.  
 PS  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 XX  
 SQ Sequence 13 AA;  
 SO  
 Query Match 89.4%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ADGPTLRWISFC 14  
 Db 1 ADGPTLRWISFC 13  
 RESULT 38  
 ID AAW35397 standard; peptide; 13 AA.  
 XX  
 AC AAW35397;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "COCCH2-alanine linked via CH2 group to Cys13"  
 FT Modified-site 13 /note= "NH2-cytosine linked via thiol group to Ala1"  
 FT  
 XX  
 PN W09640750-A1.  
 XX  
 XX 19-DEC-1996.  
 PD  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 XX 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barret RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 XX WPI; 1997-052226/05.  
 DR  
 XX Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 6; Page 63; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX Sequence 13 AA;

Query Match 89.4%; Score 76; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADGPTLRWISFC 14  
 |||||

Db 1 ADGPTLRWISFC 13

RESULT 39

AAU25997  
 ID AAU25997 standard; peptide; 13 AA.

AC AAU25997;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #183.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KM haemostatic; thrombocytopaenia; chemotherapy; radiation therapy; ELISA;  
 KM bone marrow transplantation; haematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

PN US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;  
 PI Yin Q;

DR WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopaenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 143-144; 128pp; English.

XX Sequences AAU25915-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopaenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopaenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopaenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 13 AA;

Query Match 89.4%; Score 76; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CADGPTLRWISF 13  
 |||||

Db 1 CADGPTLRWISF 13

RESULT 40

AAU25984  
 ID AAU25984 standard; peptide; 13 AA.

AC AAU25984;

DT 17-DEC-2001 (first entry)

DE Human thrombopoietin receptor (TPO-R) activator peptide #170.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KM haemostatic; thrombocytopaenia; chemotherapy; radiation therapy; ELISA;  
 KM bone marrow transplantation; haematological disorder; platelet disorder;  
 KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

PN US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

PA (GLAXO ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;  
 PI Yin Q;

DR WPI; 2001-564142/63.

PT Activating thrombopoietin receptors in cells, used to treat  
 PT thrombocytopaenia and hematological disorders, comprises contacting cells  
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 137; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopaenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopaenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopaenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 13 AA;

Query Match 89.4%; Score 76; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADGPTLRWISFC 14  
 |||||  
 DB 1 ADGPTLRWISFC 13

RESULT 41

AAW35398  
 ID AAW35398 standard; peptide; 14 AA.

AC AAW35398;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1.14

FT Modified-site 1 /note= "Homocysteine"

FT Modified-site 14 /note= "NH2-Cys"

PN MO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 6; Page 63; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX Sequence 14 AA;

Query Match 89.4%; Score 76; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADGPTLRWISFC 14  
 |||||  
 DB 2 ADGPTLRWISFC 14

RESULT 42

AAW35396  
 ID AAW35396 standard; peptide; 14 AA.

AC AAW35396;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1.14

FT Modified-site 1 /note= "penicillamine"

FT Modified-site 14 /note= "NH2-Cys"

PN MO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

PI Dower WJ, Barret RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 6; Page 63; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

OY 2 ADGPTLRWISFC 14  
 DB 2 ADGPTLRWISFC 14

## RESULT 43

AAW35402 standard; peptide, 14 AA.

AAW35402;

11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;

haematological disorder; thrombocytopenia; chemotherapy;

signal transduction; receptor activation; cell culture.

Synthetic.

Key Location/Qualifiers

Disulfide-bond 1.14

Modified-site /note= "D-form residue, Penicillamine"

Modified-site 14 /note= "NH2-D-Cys"

MO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAXO) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Mattheaetis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

peptides and peptide mimetics which bind to and activate the

thrombopoietin receptor - useful in treatment of haematological

disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Example 6; Page 64; 106pp; English.

The present peptide, which binds the thrombopoietin receptor (TR), can be

used to treat disorders which are susceptible to treatment with a

thrombopoietin agonist, preferably haematological disorders and

thrombocytopenia resulting from chemotherapy, radiation therapy or bone

marrow transfusions. It can also be used diagnostically, e.g. to

investigate the mechanism of thrombopoietin signal transduction and

receptor activation, or to maintain the proliferation and growth of

thrombopoietin dependent cell lines

Sequence 14 AA:

Query Match 89.4%; Score 76; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ADGPTLRWISFC 14

DB 2 ADGPTLRWISFC 14

## RESULT 44

AAU25987

ID AAU25987 standard; peptide, 14 AA.

AAU25987;

18-DEC-2001 (first entry)

Human thrombopoietin receptor (TPO-R) activator peptide #173.

Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

haematological disorder; chemotherapy; radiation therapy; ELISA;

bone marrow transplantation; haematological disorder; platelet disorder;

enzyme-linked immunosorbent assay; in situ staining; biological fluid;

tissue homogenate; fluorescence-activated cell sorting; Western blotting;

in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacti gene.

Homo sapiens.

US6251864-B1.

26-JUN-2001.

01-MAR-2000; 2000US-00516704.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

07-JUN-1996; 96WO-US009623.

15-AUG-1996; 96US-00699027.

(GLAXO) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;

Balasubramanian P, Wagstrom CR, Hendren RW, Deprience RB, Poddurti S;

Yin Q;

WPI; 2001-564142/63.

Activating thrombopoietin receptors in cells, used to treat

thrombocytopenia and hematological disorders, comprises contacting cells

with peptides and peptide mimetics attached to hydrophilic polymers.

Disclosure; Col 139; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

bind to and activate the human thrombopoietin receptor (TPO-R). Methods

of activating thrombopoietin receptors in cells comprise contacting the

cells with effective amounts of peptides and peptide mimetics attached to

hydrophilic polymers. The methods are used to treat thrombocytopenia such

as that due to chemotherapy, radiation therapy or bone-marrow

transplantation and to prevent thrombocytopenia in patients at risk. The

sequences are used to treat and prevent haematological disorders

including thrombocytopenia and platelet disorders. They are used in vitro

as unique tools for understanding the biological role of thrombopoietin

(TPO) and to develop other compounds that bind to and activate the TPO

receptor. The peptides can be used to detect TPO receptors on living

Sequence 14 AA:

Query Match 89.4%; Score 76; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CADGPTLRWISF 13

DB 1 CADGPTLRWISF 13

## RESULT 45

Search completed: September 1, 2005, 16:12:14  
Job time : 65.3597 secs

AAU25983  
ID AAU25983 standard; peptide; 14 AA.  
AC AAU25983;  
XX  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human thrombopoietin receptor (TPO-R) activator peptide #169.  
XX  
KM Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
KM haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
KM bone marrow transplantation; haematological disorder; platelet disorder;  
KM enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
KM tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
KM in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
XX  
OS Homo sapiens.  
XX  
PN US6251864-B1.  
XX  
PD 26-JUN-2001.  
XX  
PF 01-MAR-2000; 2000US-00516704.  
XX  
PR 07-JUN-1995; 95US-00478128.  
PR 07-JUN-1995; 95US-00485301.  
PR 07-JUN-1996; 96WO-US009623.  
PR 15-AUG-1996; 96US-00699027.  
XX  
XX  
PA (GLAXO ) GLAXO GROUP LTD.  
PI Dower WJ, Barrett RW, Cwirja SE, Gates CM, Schatz PJ,  
PI Balasubramanian P, Wagstrom CR, Hendren RM, Deprince RB, Podduturi S;  
PI Yin Q;  
XX  
XX  
DR WPI; 2001-564142/63.  
XX  
XX  
PT Activating thrombopoietin receptors in cells, used to treat  
PT thrombocytopenia and hematological disorders, comprises contacting cells  
PT with peptides and peptide mimetics attached to hydrophilic polymers.  
XX  
XX  
PS Disclosure; Col 135-137; 128pp; English.  
XX  
XX  
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
CC of activating thrombopoietin receptors in cells comprise contacting the  
CC cells with effective amounts of peptides and peptide mimetics attached to  
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
CC as that due to chemotherapy, radiation therapy or bone-marrow  
CC transplantation and to prevent thrombocytopenia in patients at risk. The  
CC sequences are used to treat and prevent haematological disorders  
CC including thrombocytopenia and platelet disorders. They are used in vitro  
CC as unique tools for understanding the biological role of thrombopoietin  
CC (TPO) and to develop other compounds that bind to and activate the TPO  
CC receptor. The peptides can be used to detect TPO receptors on living  
CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
CC in purified or natural biological materials. They may also be used for in  
CC situ staining, fluorescence-activated cell sorting, Western blotting and  
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
CC be used for in vitro expansion of megakaryocytes and their committed  
CC progenitors alone or in conjunction with additional cytokines  
XX  
XX  
SQ Sequence 14 AA;  
  
Query Match 89.4%; Score 76; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 ADGPTLRWISFC 14  
| | | | | | | | | | | | | | | |  
Db 2 ADGPTLRWISFC 14

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 15:57:33 ; Search time 10.6763 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-12

Perfect score: 85  
Sequence: 1 CADGPTLRWISFC 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	52.9	108	2 T49731	hypothetical prote
2	44	51.8	974	2 S34189	starch phosphoryla
3	44	51.8	1022	1 S00503	Na+/K+-exchanging
4	44	51.8	1023	2 A24414	Na+/K+-exchanging
5	43	50.6	245	2 T47701	translational initia
6	43	50.6	490	2 T09084	phosphatidylinosit
7	43	50.6	1010	2 B37227	Na+/K+-exchanging
8	43	50.6	1013	1 S00801	Na+/K+-exchanging
9	43	50.6	1013	2 C24639	Na+/K+-exchanging
10	43	50.6	1017	2 A37227	Na+/K+-exchanging
11	43	50.6	1020	2 A34474	Na+/K+-exchanging
12	43	50.6	1020	2 B24639	Na+/K+-exchanging
13	43	50.6	1021	1 PWSHNA	Na+/K+-exchanging
14	43	50.6	1021	1 S04630	Na+/K+-exchanging
15	43	50.6	1021	2 A28199	Na+/K+-exchanging
16	43	50.6	1021	2 B24862	Na+/K+-exchanging
17	43	50.6	1022	2 S49127	Na+/K+-exchanging
18	43	50.6	1023	1 A24639	Na+/K+-exchanging
19	43	50.6	1023	1 S24650	Na+/K+-exchanging
20	43	50.6	1025	2 A60444	Na+/K+-exchanging
21	43	50.6	1027	1 PWCNM	Na+/K+-exchanging
22	43	50.6	1038	1 S03632	Na+/K+-exchanging
23	42.5	50.0	1004	2 JH0470	Na+/K+-exchanging
24	42	49.4	522	2 D69226	hypothetical prote
25	42	49.4	522	2 S62941	hypothetical prote
26	42	49.4	725	2 A11544	conserved membrane
27	42	49.4	842	2 T12091	starch phosphoryla
28	42	49.4	842	2 S07755	hypothetical prote
29	41	48.2	189	2	

30	41	48.2	273	2 H70849	hypothetical prote
31	41	48.2	953	2 E84853	hypothetical prote
32	41	48.2	975	2 T10947	starch phosphoryla
33	41	48.2	966	2 PHE0AG	starch phosphoryla
34	41	48.2	971	2 T09210	starch phosphoryla
35	41	48.2	1000	2 S47243	starch phosphoryla
36	41	48.2	1616	2 T17884	S-layer protein -
37	40	47.1	98	2 A70301	ribosomal protein
38	40	47.1	152	2 S21826	T-cell receptor be
39	40	47.1	169	1 ICMS2	interleukin-2 prec
40	40	47.1	169	1 S37289	interleukin-2 prec
41	40	47.1	169	2 E95908	hypothetical prote
42	40	47.1	217	2 S46354	pol polyprotein -
43	40	47.1	252	2 B97072	probable hydrolase
44	40	47.1	389	2 B69096	corrinoid/iron-sul
45	40	47.1	409	2 S36113	LIS-1 protein - hu
46	40	47.1	410	2 S48052	platelet-activatin
47	40	47.1	457	2 C82720	UDP-N-acetylglucos
48	40	47.1	526	2 A86440	58.5K hypothetical
49	40	47.1	556	2 S30484	pol polyprotein -
50	40	47.1	556	2 S30483	pol polyprotein -
51	40	47.1	838	1 A40995	starch phosphoryla
52	40	47.1	1034	1 GNLJCA	HIV-1 retropepsin
53	40	47.1	1035	1 GNLJGG	HIV-1 retropepsin
54	40	47.1	1036	1 GNLJG2	HIV-1 retropepsin
55	40	47.1	1055	1 GNLJST	HIV-1 retropepsin
56	40	47.1	1055	1 S53092	pol polyprotein -
57	40	47.1	1733	1 RNBRY2L	DNA-directed RNA p
58	40	47.1	3083	2 AH2493	hypothetical prote
59	39	45.9	113	2 D72595	hypothetical prote
60	39	45.9	180	2 T44944	hypothetical prote
61	39	45.9	207	2 B75327	hypothetical prote
62	39	45.9	331	2 B48445	glyceralddehyde-3-p
63	39	45.9	361	2 F91207	hypothetical prote
64	39	45.9	361	2 H86053	hypothetical prote
65	39	45.9	361	2 C65171	hypothetical 41.0
66	39	45.9	379	2 I48133	ubiquinol-cytochro
67	39	45.9	379	2 I48132	ubiquinol-cytochro
68	39	45.9	379	2 I48134	ubiquinol-cytochro
69	39	45.9	379	2 I48180	ubiquinol-cytochro
70	39	45.9	428	2 JH0634	site-specific DNA-
71	39	45.9	491	2 F83383	probable flavin-bi
72	39	45.9	534	2 S69714	hypothetical prote
73	39	45.9	566	2 B84271	glutemyl-tRNA synt
74	39	45.9	591	2 S54788	calcium-stimulated
75	39	45.9	789	2 S28259	androgen-regulated
76	39	45.9	817	2 A82511	glycogen phosphory
77	39	45.9	942	2 A12530	hypothetical prote
78	39	45.9	1008	2 S38003	translation elonga
79	38.5	45.3	505	2 T19971	hypothetical prote
80	38.5	45.3	506	2 T19973	hypothetical prote
81	38	44.7	56	2 T03658	phosphoenolpyruvat
82	38	44.7	142	2 AF0961	heat shock protein
83	38	44.7	154	2 AC0496	heat shock protein
84	38	44.7	252	2 C84522	22 kDa peroxisomal
85	38	44.7	266	2 E90354	hypothetical prote
86	38	44.7	410	1 DBPSXA	3-methyl-2-oxobuta
87	38	44.7	410	2 C83365	2-oxoisovalerate d
88	38	44.7	477	2 T25798	hypothetical prote
89	38	44.7	480	2 H84747	probable steroid d
90	38	44.7	511	2 D70522	probable papai pro
91	38	44.7	566	2 T09154	glucose-6-phosphat
92	38	44.7	568	2 S57830	glucose-6-phosphat
93	38	44.7	569	2 S23542	glucose-6-phosphat
94	38	44.7	569	2 S41806	glucose-6-phosphat
95	38	44.7	569	2 S57831	conserved hypochet
96	38	44.7	725	1 AB1187	replication licens
97	38	44.7	770	1 T03920	conserved hypochet
98	38	44.7	777	1 G69773	starch phosphoryla
99	38	44.7	841	2 T45633	starch phosphoryla
100	38	44.7	886	1 JCS085	replication licens

## ALIGNMENTS

```
RESULT 1
T49731
hypothetical protein B24B19.30 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #ext_change 18-Aug-2000
C:Accession: T49731
R:Schulte, U.; Aigm, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49731
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <SCH>
A:Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.30
A:Experimental source: BAC clone B24B19; strain OR74A
C:Genetics:
A:Gene: NCSP:B24B19.30
A:Map position: 6
C:Superfamily: Neurospora crassa hypothetical protein B24B19.30

Query Match          52.9%; Score 45; DB 2; Length 108;
Best Local Similarity 50.0%; Pred. No. 3.6;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      1 CADGPTLRWISFC 14
      | | | | | | | |
Db      70 CQCQPIRWLWLMWC 83

RESULT 2
S34189
starch phosphorylase (EC 2.4.1.1) L - potato
C:Species: Solanum tuberosum (potato)
C:Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #ext_change 09-Jul-2004
C:Accession: S53489; S34189
R:Sommerwald, U.; Basner, A.; Greve, B.; Steup, M.
Plant Mol. Biol. 27, 567-576, 1995
A>Title: A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inh
A:Reference number: S53489; MUID:95201249; PMID:7694019
A:Accession: S53489
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-974 <SO2>
A:Cross-references: UNIPROT:P53535; EMBL:X73684; NID:G313348; PIDN:CAA52036.1; PID:G3133
C:Superfamily: glucan phosphorylase
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphate
F:820/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match          51.8%; Score 44; DB 2; Length 974;
Best Local Similarity 58.3%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      3 DGPTLRWISFC 14
      | | | | | | | |
Db      619 NGVTPRRWISFC 630

RESULT 3
S00503
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - Pacific electric ray
C:Species: Torpedo californica (Pacific electric ray)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #ext_change 09-Jul-2004
C:Accession: S00503; S28885; S29880
R:Kawakami, K.; Noguchi, S.; Noda, M.; Takahashi, H.; Ohta, T.; Kawamura, M.; Nojima, H.
Nature 316, 733-736, 1985
A>Title: Primary structure of the alpha-subunit of Torpedo californica Na(+)+K(+)ATPase
A:Reference number: S00503; MUID:85296307; PMID:2993905
A:Accession: S00503
A:Molecule type: mRNA
A:Residues: 1-1022 <KAW1>
```

```
A:Cross-references: UNIPROT:P05025; EMBL:X02810; NID:G64399; PIDN:CAA26578.1; PID:G64400
A:Accession: S28885
A:Molecule type: protein
A:Residues: 228-240/431-438/535-550/671-690/1011-1022 <KAW2>
R:Ohta, T.; Nagano, K.; Yoshida, M.
Proc. Natl. Acad. Sci. U.S.A. 83, 2071-2075, 1986
A>Title: The active site structure of Na(+)/K(+)-transporting ATPase: location of the 5'
A:Reference number: S29880; MUID:86177549; PMID:3008150
A:Accession: S29880
A:Molecule type: protein
A:Residues: 386-402/502-512/671-689/887-906 <OHT>
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp
F:96-120/Domain: transmembrane #status predicted <TM1>
F:130-149/Domain: transmembrane #status predicted <TM2>
F:150-290/Domain: intracellular #status predicted <INT2>
F:291-313/Domain: transmembrane #status predicted <TM3>
F:320-348/Domain: transmembrane #status predicted <TM4>
F:349-785/Domain: intracellular #status predicted <INT3>
F:586-782/Domain: ATPase nucleotide-binding domain homology <ATN>
F:786-809/Domain: transmembrane #status predicted <TM5>
F:848-873/Domain: transmembrane #status predicted <TM6>
F:874-951/Domain: intracellular #status predicted <INT4>
F:952-977/Domain: transmembrane #status predicted <TM7>
F:978-1022/Domain: extracellular #status predicted <EXT>
F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:507/Binding site: ATP (Lys) #status predicted
F:176/720/725/Active site: Asp, Asp, Lys #status predicted

Query Match          51.8%; Score 44; DB 1; Length 1022;
Best Local Similarity 70.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 PTLRWISFC 14
      | | | | | | | |
Db      84 PTPRWIKFC 93

RESULT 4
A24414
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - human
N:Alternate names: sodium pump; sodium/potassium transporting ATPase alpha-A chain
C:Species: Homo sapiens (man)
C:Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #ext_change 09-Jul-2004
C:Accession: A24414; A27795; A39910; 160116; S09171
R:Kawakami, K.; Ohta, T.; Nojima, H.; Nagano, K.
J. Biochem. 100, 389-397, 1986
A>Title: Primary structure of the alpha-subunit of human Na,K-ATPase deduced from cDNA s
A:Reference number: A24414; MUID:87057096; PMID:2430951
A:Accession: A24414
A:Molecule type: mRNA
A:Residues: 1-1023 <KAW>
A:Cross-references: UNIPROT:P05023; EMBL:X04297; NID:G28926; PIDN:CAA27840.1; PID:G28927
R:Shull, M.M.; Lingrel, J.B.
Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987
A>Title: Multiple genes encode the human Na+,K+-ATPase catalytic subunit.
A:Reference number: A94158; MUID:87231946; PMID:3035563
A:Accession: A27795
A:Molecule type: DNA
A:Residues: 168-189/213-214/'X', 216-244 <SHU>
R:Chehab, F.F.; Kan, Y.W.; Law, M.L.; Hartz, J.; Kao, F.T.; Blotstein, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 7901-7905, 1987
A>Title: Human placental Na+,K+-ATPase alpha subunit: cDNA cloning, tissue expression, E
A:Reference number: A39910; MUID:88068506; PMID:2891135
A:Accession: A39910
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 199-942 <CHE>
A:Cross-references: GB:J035007
R:Shull, M.M.; Pugh, D.G.; Lingrel, J.B.
Genomics 6, 451-460, 1990
A>Title: The human Na,K-ATPase alpha 1 gene: characterization of the 5'-flanking region
A:Reference number: I60116; MUID:90228961; PMID:1970326
```



A:Accession: I60116  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
C:Species: Arabidopsis thaliana  
A:Molecule type: DNA  
A:Residues: 1-61 <RBS>  
A:Cross-references: GB:M30310; NID:g179206; PIDN:AAAS1801.1; PID:g179208  
C:Genetics:  
A:Gene: GDB:ATP1A1  
A:Cross-references: GDB:119711; OMIM:182310  
A:Map position: 1p13-1p11  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; heterodimer; hydrolase; ion transport; osmoregulation; phosphoprotein;  
F:6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MNT>  
F:6-95/Domain: intracellular #status predicted <INT1>  
F:96-120/Domain: transmembrane #status predicted <TM1>  
F:130-149/Domain: transmembrane #status predicted <TM2>  
F:150-290/Domain: intracellular #status predicted <INT2>  
F:291-313/Domain: transmembrane #status predicted <TM3>  
F:320-348/Domain: transmembrane #status predicted <TM4>  
F:349-786/Domain: intracellular #status predicted <INT3>  
F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:787-810/Domain: transmembrane #status predicted <TM5>  
F:849-874/Domain: transmembrane #status predicted <TM6>  
F:875-952/Domain: intracellular #status predicted <INT4>  
F:953-978/Domain: transmembrane #status predicted <TM7>  
F:979-1023/Domain: extracellular #status predicted <EXT>  
F:1576/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:508/Binding site: ATP (lys) #status predicted  
F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 51.8%; Score 44; DB 2; Length 1023;  
Best Local Similarity 70.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14  
DB 84 PTLREWIKFC 93

RESULT 5  
T47701  
translation initiation factor eIF-6-like protein [imported] - Arabidopsis thaliana  
M:Alternate names: protein P116.30  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47701  
R:Bens, V.; Murnbach, E.; Dronek, H.; Anserge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.R.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24473  
A:Accession: T47701  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-245 <BEN>  
A:Cross-references: UNIPROT:Q9M060; EMBL:AL161667  
A:Experimental source: cultivar Columbia; BAC clone P116  
C:Genetics:  
A:Map position: 3  
A:Introns: 4/1; 36/2; 65/1; 80/1; 123/3; 160/3  
A:Note: P116.30  
C:Superfamily: conserved hypothetical protein YPR016c

Query Match 50.6%; Score 43; DB 2; Length 245;  
Best Local Similarity 53.8%; Pred. No. 17;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ADGPTLRWISFC 14  
DB 194 AAGTVDWTSFC 206

RESULT 6  
T09084  
phosphatidylinositol 3-kinase - Chlamydomonas reinhardtii (fragment)  
C:Species: Chlamydomonas reinhardtii

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09084  
R:Molendijk, A.J.; Irvine, R.F.  
Plant Mol. Biol. 37, 53-66, 1998  
A:Title: Inositolide signalling in Chlamydomonas: Characterization of a phosphatidylinositol  
A:Reference number: Z16411; MUID:98281574; PMID:9620264  
A:Accession: T09084  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-430 <MOI>  
A:Cross-references: UNIPROT:O04270; EMBL:U97663; NID:g2109290; PIDN:AAAS0018.1; PID:g21  
A:Experimental source: strain cw-15  
C:Genetics:  
A:Introns: 265/3; 331/3; 370/3; 455/1; 481/3

Query Match 50.6%; Score 43; DB 2; Length 490;  
Best Local Similarity 57.1%; Pred. No. 34;  
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 3 DGPTRLR--EWISFC 14  
DB 250 DGPSTARWDEWLTFC 263

RESULT 7  
B37227  
Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha-3 chain - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 09-Jul-2004  
C:Accession: B37227; 150395  
R:Takeyasu, K.; Lemas, V.; Fambrough, D.M.  
Am. J. Physiol. 259, C619-C630, 1990  
A:Title: Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.  
A:Reference number: A37227; MUID:91023019; PMID:2171348  
A:Accession: B37227  
A:Molecule type: mRNA  
A:Residues: 1-1010 <TA2>  
A:Cross-references: UNIPROT:P24798; GB:M59960; NID:g212407; PIDN:AA48982.1; PID:g21240  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; sodium c  
F:574-770/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:202-470/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:363/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:495/Binding site: ATP (lys) #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1010;  
Best Local Similarity 60.0%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14  
DB 71 PTLREWIKFC 80

RESULT 8  
S00801  
Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha-3 chain - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
C:Accession: S00801; S04019; A27397; S02275  
R:Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Broude, N.E.; Ushkarev, Y.A.; Melkov, A.M.;  
dyanov, N.N.; Sverdlov, E.D.  
FEBS Lett. 233, 87-94, 1988  
A:Title: Family of human Na,K-ATPase genes. Structure of the gene for the catalytic sub  
A:Reference number: S00801; MUID:8825304; PMID:2838329  
A:Accession: S00801  
A:Molecule type: DNA  
A:Residues: 1-1013 <OV>  
A:Cross-references: UNIPROT:P13637; EMBL:M57456  
R:Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkarev, Y.A.; Melkov, A.M.; Smir  
ov, N.N.; Ovchinnikov, Y.A.  
Dokl. Biochem. 297, 426-431, 1987  
A:Title: Family of human Na(+),K(+)-ATPase genes. Structure of the gene of isoform alph

```

A:Reference number: S04019
A:Accession: S04019
A:Molecule type: DNA
A:Residues: 1, 'E1H', 3-1013 <SVE1>
A:Cross-references: EMBL:X12910; NID:q28963
A:Note: the authors translated the codon TTC for residue 283 as Ser and TCT for residue
A:Note: this paper is a translation of the Russian paper published in Dokl. Akad. Nauk S
R.Sverdlov, E.D.; Monastyrskaya, G.S.; Broide, N.E.; Ushakov, Y.A.; Alilmet, R.L.; M
tina, M.B.; Sverdlov, V.B.; Modyanov, N.N.; Ovchinnikov, Y.A.
FEBS Lett. 217, 275-278, 1987
A:Title: The family of human Na+/K+-ATPase genes. No less than five genes and/or pseudog
A:Reference number: A27397; MUID:87247232; PMID:3036582
A:Accession: A27397
A:Molecule type: mRNA
A:Residues: 243-434 <SVE2>
A:Cross-references: GB:M27570
C:Genetics:
A:Gene: GDB:ATPIA3
A:Cross-references: GDB:119713; OMIM:182350
A:Map position: 19q13.2-19q13.2
A:Introns: 2/3; 31/3; 51/3; 119/3; 157/3; 202/3; 242/1; 331/3; 398/1; 435/2; 479/3; 544/
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F:86-110/Domain: transmembrane #status predicted <TM1>
F:120-138/Domain: transmembrane #status predicted <TM2>
F:140-280/Domain: transmembrane #status predicted <TM3>
F:281-303/Domain: transmembrane #status predicted <TM4>
F:310-333/Domain: transmembrane #status predicted <TM5>
F:339-776/Domain: intracellular #status predicted <INT3>
F:577-773/Domain: intracellular #status predicted <INT4>
F:777-800/Domain: transmembrane #status predicted <TM5>
F:839-864/Domain: transmembrane #status predicted <TM6>
F:865-942/Domain: intracellular #status predicted <INT4>
F:943-966/Domain: transmembrane #status predicted <TM7>
F:969-1013/Domain: extracellular #status predicted <EXT>
F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:498/Binding site: Asp (Lys) #status predicted
F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match          50.6%; Score 43; DB 1; Length 1013;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      5 PTLREWISFC 14
DB      74 PTPPEWVKFC 83

RESULT 9
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - rat
N:Alternate names: Na+/K+-transporting ATPase alpha(III) chain
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1988 #sequence revision 23-Apr-1993 #text_change 09-Jul-2004
C:Accession: C24639; S00514; B27180; A60470
R:Shull, G.E.; Greeb, U.; Lingrel, J.B.
Biochemistry 25, 8125-8132, 1986
A:Title: Molecular cloning of three distinct forms of the Na+/K+-ATPase alpha-subunit fr
A:Reference number: A90512; MUID:87128908; PMID:3028470
A:Accession: C24639
A:Molecule type: mRNA
A:Residues: 1-1013 <SHU>
A:Cross-references: UNIPROT:P06687; EMBL:M14513; NID:g203030; PID:AAA40777.1; PID:g2030
A:Note: in the authors' translation 405-Ser is shown after residue 409 and, consequently
R.Harz, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.; N
J Biochem 102, 43-58, 1987
A:Title: Primary structures of two types of alpha-subunit of rat brain Na(+) ,K(+) -ATPase
A:Reference number: S00460; MUID:88032933; PMID:2822682
A:Accession: S00514
A:Molecule type: mRNA
A:Residues: 1-907, 'C', 909-1013 <HAR>
A:Cross-references: EMBL:X05883; NID:g55769; PID:CAA29307.1; PID:g55770
R.Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.

```

```

J Cell Biol. 105, 1855-1865, 1987
A:Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural a
A:Reference number: A92749; MUID:88033255; PMID:2822726
A:Accession: B27180
A:Molecule type: mRNA
A:Residues: 1, 'NL', 4-103, 'R', 105-113, 'E', 115-127, 'G', 129-148, 'Q', 150-151, 'T', 153-165, 'D'
A:Cross-references: EMBL:M26649; NID:g205633; PID:AAA1672.1; PID:g205634
A:Note: the authors translated the codon CAG for residue 149 as Glu, GCG for residue 194
R.Hsu, Y.M.; Guidotti, G.
Biochemistry 28, 569-573, 1989
A:Title: Rat brain has the alpha3 form of the (Na+/K+)ATPase.
A:Reference number: A60470; MUID:89229049; PMID:2540801
A:Accession: A60470
A:Molecule type: protein
A:Residues: 117-132;586-595, 'X', 597-601 <HSU>
A:Comment: The alpha-3 form appears to be highly ouabain-inhibitable, as is alpha-2 but
C:Genetics:
A:Gene: NKAA3
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F:86-110/Domain: transmembrane #status predicted <TM1>
F:120-138/Domain: transmembrane #status predicted <TM2>
F:140-280/Domain: intracellular #status predicted <INT2>
F:281-303/Domain: transmembrane #status predicted <TM3>
F:310-338/Domain: transmembrane #status predicted <TM4>
F:339-776/Domain: intracellular #status predicted <INT3>
F:577-773/Domain: intracellular #status predicted <INT4>
F:777-800/Domain: transmembrane #status predicted <TM5>
F:839-864/Domain: transmembrane #status predicted <TM6>
F:865-942/Domain: intracellular #status predicted <INT4>
F:943-966/Domain: transmembrane #status predicted <TM7>
F:969-1013/Domain: extracellular #status predicted <EXT>
F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:498/Binding site: Asp (Lys) #status predicted
F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match          50.6%; Score 43; DB 2; Length 1013;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      5 PTLREWISFC 14
DB      74 PTPPEWVKFC 83

RESULT 10
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - chicken
C:Species: Gallus gallus (chicken)
C:Date: 16-Sep-1992 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: I50394; A37227
R:Takeyasu, K.; Lemas, M.; Fambrough, D.M.
Am. J. Physiol. 259, C619-C630, 1990
A:Title: Stability of the Na+/K+-ATPase alpha-subunit isoforms in evolution.
A:Reference number: A37227; MUID:91023019; PMID:2171348
A:Accession: A37227
A:Molecule type: mRNA
A:Residues: 3-1017 <TA2>
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F:86-110/Domain: transmembrane #status predicted <TM1>
F:120-138/Domain: transmembrane #status predicted <TM2>
F:140-280/Domain: intracellular #status predicted <INT2>
F:281-303/Domain: transmembrane #status predicted <TM3>
F:310-338/Domain: transmembrane #status predicted <TM4>
F:339-776/Domain: intracellular #status predicted <INT3>
F:577-773/Domain: intracellular #status predicted <INT4>
F:777-800/Domain: transmembrane #status predicted <TM5>
F:839-864/Domain: transmembrane #status predicted <TM6>
F:865-942/Domain: intracellular #status predicted <INT4>
F:943-966/Domain: transmembrane #status predicted <TM7>
F:969-1013/Domain: extracellular #status predicted <EXT>
F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:498/Binding site: Asp (Lys) #status predicted
F:707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match          50.6%; Score 43; DB 2; Length 1013;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      5 PTLREWISFC 14
DB      74 PTPPEWVKFC 83

RESULT 10
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - chicken
C:Species: Gallus gallus (chicken)
C:Date: 16-Sep-1992 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: I50394; A37227
R:Takeyasu, K.; Lemas, M.; Fambrough, D.M.
Am. J. Physiol. 259, C619-C630, 1990
A:Title: Stability of the Na+/K+-ATPase alpha-subunit isoforms in evolution.
A:Reference number: A37227; MUID:91023019; PMID:2171348
A:Accession: A37227
A:Molecule type: mRNA
A:Residues: 3-1017 <TA2>
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F:86-110/Domain: transmembrane #status predicted <TM1>
F:120-138/Domain: transmembrane #status predicted <TM2>
F:140-280/Domain: intracellular #status predicted <INT2>
F:281-303/Domain: transmembrane #status predicted <TM3>
F:310-338/Domain: transmembrane #status predicted <TM4>
F:339-776/Domain: intracellular #status predicted <INT3>
F:577-773/Domain: intracellular #status predicted <INT4>
F:777-800/Domain: transmembrane #status predicted <TM5>
F:839-864/Domain: transmembrane #status predicted <TM6>
F:865-942/Domain: intracellular #status predicted <INT4>
F:943-966/Domain: transmembrane #status predicted <TM7>
F:969-1013/Domain: extracellular #status predicted <EXT>
F:366/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:498/Binding site: Asp (Lys) #status predicted
F:707,711,716/Active site: Asp, Asp, Lys #status predicted

```

Query Match 50.6%; Score 43; DB 2; Length 1017;  
 Best Local Similarity 60.0%; Pred. No. 64;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 DB 79 PTPPEWVKFC 88

## RESULT 11

A34474  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - human  
 N/Alternate names: Na+/K+-exchanging ATPase alpha chain-4; sodium/potassium transporting  
 C/Species: Homo sapiens (man)  
 C/Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 09-Jul-2004  
 C/Accession: A34474; B27795; D27397  
 R/Shull, M.M.; Pugh, D.G.; Lingrel, J.B.  
 J. Biol. Chem. 264, 17532-17543, 1989  
 A/Title: Characterization of the human Na,K-ATPase alpha2 gene and identification of int  
 A/Reference number: A34474; MUID:90008924; PMID:2477373  
 A/Accession: A34474  
 A/Molecule type: DNA  
 A/Residues: 1-1020 <SHU>  
 A/Cross-references: UNIPROT:P50993; GB:J05096; NID:g179164; PIDN:AAA51797.1; PID:g179165  
 R/Shull, M.M.; Lingrel, J.B.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
 A/Title: Multiple genes encode the human Na+,K+-ATPase catalytic subunit.  
 A/Reference number: A94158; MUID:87231946; PMID:3035563  
 A/Accession: B27795  
 A/Molecule type: DNA  
 A/Residues: 211-249 <SH2>  
 A/Cross-references: GB:M16795; NID:g179196; PIDN:AAA51799.1; PID:g553194  
 R/Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Alilmetz, R.L.; W  
 PMSA Lett. 217, 275-278, 1987  
 A/Title: The family of human Na+,K+-ATPase genes. No less than five genes and/or pseudog  
 A/Reference number: A27397; MUID:87247232; PMID:3036582  
 A/Accession: D27397  
 A/Molecule type: DNA  
 A/Residues: 251-442 <SVE>  
 A/Cross-references: GB:M27571  
 C/Genetics:  
 A/Gene: GDB:ATP1A2  
 A/Cross-references: GDB:119712; OMIM:182340  
 A/Map position: 1q21-1q23  
 C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C/Keyword: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F/6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MAT>  
 F/6-93/Domain: intracellular #status predicted <INT1>  
 F/94-118/Domain: transmembrane #status predicted <TM1>  
 F/128-147/Domain: transmembrane #status predicted <TM2>  
 F/148-288/Domain: intracellular #status predicted <INT2>  
 F/189-311/Domain: transmembrane #status predicted <TM3>  
 F/318-346/Domain: transmembrane #status predicted <INT3>  
 F/347-783/Domain: intracellular #status predicted <INT4>  
 F/584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F/846-807/Domain: transmembrane #status predicted <TM5>  
 F/872-949/Domain: intracellular #status predicted <INT4>  
 F/950-975/Domain: transmembrane #status predicted <TM7>  
 F/976-1020/Domain: extracellular #status predicted <EXT>  
 F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F/505/Binding site: ATP (lys) #status predicted  
 F/714,718,723/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1020;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 DB 82 PTPPEWVKFC 91

## RESULT 12

B24639  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - rat

N/Alternate names: Na+/K+-transporting ATPase alpha-plus chain  
 C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
 C/Accession: B24639

R/Shull, G.E.; Greed, J.; Lingrel, J.B.  
 Biochemistry 25, 8125-8132, 1986

A/Title: Molecular cloning of three distinct forms of the Na+,K+-ATPase alpha-subunit f  
 A/Reference number: A90512; MUID:87128908; PMID:3028470

A/Accession: B24639  
 A/Molecule type: mRNA

A/Residues: 1-1020 <SHU>  
 A/Cross-references: UNIPROT:P06686; EMBL:M14512; NID:g203028; PIDN:AAA40776.1; PID:g203

C/Genetics:  
 A/Gene: NKA2

C/Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C/Keyword: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans

F/6-1020/Product: Na+/K+-transporting ATPase alpha-2 chain #status predicted <MAT>  
 F/6-93/Domain: intracellular #status predicted <INT1>

F/94-119/Domain: transmembrane #status predicted <TM1>  
 F/128-147/Domain: transmembrane #status predicted <INT2>

F/148-288/Domain: intracellular #status predicted <INT3>  
 F/189-311/Domain: transmembrane #status predicted <TM3>

F/318-346/Domain: transmembrane #status predicted <INT4>  
 F/347-783/Domain: intracellular #status predicted <INT3>

F/584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F/846-807/Domain: transmembrane #status predicted <TM5>

F/872-949/Domain: intracellular #status predicted <INT4>  
 F/950-975/Domain: transmembrane #status predicted <TM7>

F/976-1020/Domain: extracellular #status predicted <EXT>  
 F/374/Active site: Asp (aspartylphosphate intermediate) #status predicted

F/505/Binding site: ATP (lys) #status predicted  
 F/714,718,723/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1020;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 DB 82 PTPPEWVKFC 91

## RESULT 13

PMSNA  
 Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain precursor - sheep

N/Alternate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha cha  
 C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004  
 C/Accession: A01074; A35426

R/Shull, G.E.; Schwartz, A.; Lingrel, J.B.  
 Nature 315, 691-695, 1985

A/Title: Amino-acid sequence of the catalytic subunit of the (Na(+)+K(+)) ATPase deduce  
 A/Reference number: A01074; MUID:85296299; PMID:2993903

A/Accession: A01074  
 A/Molecule type: mRNA

A/Residues: 1-1021 <SHU>  
 A/Cross-references: UNIPROT:P04074; GB:X02813; NID:g1205; PIDN:CAA26581.1; PID:g1206

R/Hinz, H.R.; Kitley, T.L.  
 J. Biol. Chem. 265, 10260-10265, 1990

A/Title: Lysine 480 is an essential residue in the putative ATP site of lamb kidney (Na  
 A/Reference number: A35426; MUID:90285144; PMID:2162343

A/Accession: A35426  
 A/Molecule type: protein

A/Status: preliminary  
 A/Molecule type: protein

A/Residues: 475-492 <HIN>  
 C/Comment: This is the catalytic component of the active enzyme, which catalyzes the hy

drates the electrochemical gradient of sodium and potassium, providing the energy for a  
 n function.

C.Comment: This enzyme is specifically inhibited by cardiac glycosides such as digoxin  
 C.Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C.Date: 21-Sep-1998 #sequence\_revision 21-Sep-1998 #text\_change 09-Jul-2004  
 C.Accession: A28199  
 C.Keywords: ATP; hydrolase; phosphoprotein; potassium transport; sodium transport; trans  
 F.6-1021/Product: Na+/K+-transporting ATPase alpha chain #status predicted <MAT>  
 F.128-114/Domain: transmembrane #status predicted <TM1>  
 F.128-115/Domain: transmembrane #status predicted <TM2>  
 F.128-311/Domain: transmembrane #status predicted <TM3>  
 F.128-311/Domain: transmembrane #status predicted <TM4>  
 F.318-346/Domain: transmembrane #status predicted <TM5>  
 F.585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F.785-808/Domain: transmembrane #status predicted <TM6>  
 F.847-872/Domain: transmembrane #status predicted <TM7>  
 F.951-976/Domain: transmembrane #status predicted <TM8>  
 F.315/Binding site: cardiac glycoside (TTP) #status predicted  
 F.374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F.506/Binding site: ATP (Lys) #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1021;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWSFC 14  
 ||| |||  
 Db 82 PTPPEWVFC 91

## RESULT 14

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - horse  
 S04630  
 C.Species: Equus caballus (domestic horse)  
 C.Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
 C.Accession: S04630  
 R.Kano, I.; Nagai, F.; Satoh, K.; Ushiyama, K.; Nakao, T.; Kano, K.  
 FBS Lett. 250, 91-98, 1989  
 A.Title: Structure of the alpha(1) subunit of horse Na,K-ATPase gene.  
 A.Reference number: S04630; MUID:89290042; PMID:2544461  
 A.Accession: S04630  
 A.Molecule type: DNA  
 A.Residues: 1-1021 <KAN>  
 A.Cross-references: UNIPROT:P18907; EMBL:X16773; NID:g1010; PIDN:CAA34716.1; PID:8871026  
 C.Genetics:  
 A.Introns: 4/3; 39/3; 59/3; 127/3; 165/3; 210/3; 250/1; 339/3; 406/1; 442/3; 487/3; 552/3  
 C.Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C.Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F.6-1021/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>  
 F.6-93/Domain: intracellular #status predicted <INT1>  
 F.94-118/Domain: transmembrane #status predicted <TM1>  
 F.128-147/Domain: transmembrane #status predicted <TM2>  
 F.128-311/Domain: intracellular #status predicted <INT2>  
 F.128-311/Domain: transmembrane #status predicted <TM3>  
 F.318-346/Domain: transmembrane #status predicted <TM4>  
 F.318-346/Domain: transmembrane #status predicted <TM5>  
 F.347-784/Domain: intracellular #status predicted <INT3>  
 F.585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F.785-808/Domain: transmembrane #status predicted <TM6>  
 F.847-872/Domain: transmembrane #status predicted <TM7>  
 F.873-950/Domain: intracellular #status predicted <INT4>  
 F.951-976/Domain: transmembrane #status predicted <TM8>  
 F.977-1021/Domain: extracellular #status predicted <EXT>  
 F.374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F.506/Binding site: ATP (Lys) #status predicted  
 F.715-719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1021;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWSFC 14  
 ||| |||  
 Db 82 PTPPEWVFC 91

## RESULT 15

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - chicken  
 A28199

C.Species: Gallus gallus (chicken)  
 C.Date: 21-Sep-1998 #sequence\_revision 21-Sep-1998 #text\_change 09-Jul-2004  
 C.Accession: A28199  
 R.Takeyasu, K.; Tamkun, M.M.; Renaud, K.D.; Fambrough, D.M.  
 J. Biol. Chem. 263, 4347-4354, 1988  
 A.Title: Ouabain-sensitive (Na<sup>+</sup>) + K<sup>+</sup>)-ATPase activity expressed in mouse L cells by  
 A.Reference number: A28199; MUID:88153759; PMID:2831227  
 A.Accession: A28199  
 A.Status: preliminary; not compared with conceptual translation  
 A.Molecule type: mRNA  
 A.Residues: 1-1021 <TAK>  
 A.Cross-references: UNIPROT:P09572; GB:J03230; NID:g211219; PIDN:AAA48607.1; PID:g211220  
 C.Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C.Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein  
 F.585-781/Domain: transmembrane #status predicted <TM6>  
 F.213/481/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F.374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F.506/Binding site: ATP (Lys) #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1021;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWSFC 14  
 ||| |||  
 Db 82 PTPPEWVFC 91

## RESULT 16

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - pig  
 B24862  
 N.Altemate names: sodium pump alpha chain; sodium/potassium-dependent ATPase alpha chain  
 C.Species: Sus scrofa domestica (domestic pig)  
 C.Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
 C.Accession: B24862; 146572; A35504; S00011; S00502; S02569; S29762  
 R.Ovchinnikov, Y.A.; Modyanov, N.N.; Browde, N.E.; Petrunkin, K.E.; Grishin, A.V.; Arzam  
 FBS Lett. 201, 237-245, 1986  
 A.Title: Pig kidney Na<sup>+</sup>,K<sup>+</sup>-ATPase. Primary structure and spatial organization.  
 A.Reference number: A31361; MUID:86220813; PMID:2423371  
 A.Accession: B24862  
 A.Molecule type: mRNA  
 A.Residues: 1-1021 <OVCA2>  
 A.Cross-references: UNIPROT:P05024; EMBL:X03938; NID:g1897; PIDN:CAA27576.1; PID:g1898  
 A.Note: the authors translated the codon TCC for residue 391 as Phe, TCG for residue 723  
 A.Note: part of this sequence, including the amino and carboxyl end of the mature protei  
 R.Ovchinnikov, Y.A.; Modyanov, N.N.; Browde, N.E.; Petrunkin, K.E.;  
 Dokl. Biochem. 283, 270-272, 1985  
 A.Title: Amino acid sequence of the 17-kilodalton fragment of the cytoplasmic region of  
 A.Reference number: 146572  
 A.Accession: 146572  
 A.Status: preliminary; translated from GB/EMBL/DBJ  
 A.Molecule type: mRNA  
 A.Residues: 469-617 <OVCA1>  
 A.Cross-references: GB:M2512; NID:g164385; PIDN:AAA31004.1; PID:g164386  
 R.Karlisch, S.J.D.; Goldshleger, R.; Stein, W.D.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 4566-4570, 1990  
 A.Title: A 19-kDa C-terminal tryptic fragment of the alpha chain of Na/K-ATPase is essen  
 A.Reference number: A35504; MUID:90280416; PMID:2162048  
 A.Accession: A35504  
 A.Molecule type: protein  
 A.Residues: 836-845, 'R', 847-851 <KAR>  
 R.Ovchinnikov, Y.A.; Arzamazova, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Aldanova, N.  
 FBS Lett. 217, 269-274, 1987  
 A.Title: Detailed structural analysis of exposed domains of membrane-bound Na<sup>+</sup>,K<sup>+</sup>-ATPase  
 A.Reference number: S00011; MUID:87247231; PMID:3036581  
 A.Contents: annotation; membrane topology  
 R.Ovchinnikov, Y.A.; Luneva, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Arzamazova, N.M.  
 FBS Lett. 227, 230-234, 1988  
 A.Title: Topology of Na, K-ATPase: identification of the extra- and intracellular hydrox  
 A.Reference number: S02569; MUID:88112252; PMID:2448169  
 A.Contents: annotation; membrane topology  
 C.Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C.Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans

F:6-1021/Product: Na+/K+-transporting ATPase alpha chain #status experimental <MAT>  
 F:6-93/Domain: intracellular #status predicted <INT1>  
 F:194-118/Domain: transmembrane #status predicted <TM2>  
 F:128-147/Domain: transmembrane #status predicted <TM2>  
 F:148-288/Domain: intracellular #status predicted <INT2>  
 F:289-311/Domain: transmembrane #status predicted <TM3>  
 F:318-346/Domain: transmembrane #status predicted <TM4>  
 F:347-784/Domain: intracellular #status predicted <INT3>  
 F:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:785-808/Domain: transmembrane #status predicted <TM5>  
 F:847-872/Domain: transmembrane #status predicted <TM6>  
 F:873-950/Domain: intracellular #status predicted <INT4>  
 F:951-976/Domain: transmembrane #status predicted <TM7>  
 F:977-1021/Domain: extracellular #status predicted <EXT>  
 F:506/Binding site: Asp (aspartylphosphate intermediate) #status predicted  
 F:715,719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1021;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 DB 82 PTPPEWVKFC 91

## RESULT 17

S49127

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - European eel

C:Species: *Anguilla anguilla* (European eel)

C:date: 01-Feb-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004

C:Accession: S49127

R:Cutler, C.; Sanders, I.L.; Cramb, G.

submitted to the EMBL Data Library, November 1993

A:Reference number: S45093

A:Accession: S49127

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1022 &lt;CUT&gt;

C:Cross-references: UNIPROT:Q92030; EMBL:X76108; NID:9509405; PIDN:CAA53714.1; PID:95094

C:Keywords: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; transmembrane

F:586-782/Domain: ATPase nucleotide-binding domain homology &lt;ATN&gt;

F:214,482/Binding site: carbonylate (Asn) (covalent) #status predicted

F:375/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:507/Binding site: ATP (Lys) #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1022;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 DB 83 PTPPEWVKFC 92

## RESULT 18

A24639

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain [validated] - rat

N:Alternate names: Na+/K+-transporting ATPase alpha chain, kidney-type

N:Conting: Na+/K+-transporting ATPase alpha-S chain

C:Species: *Rattus norvegicus* (Norway rat)

C:date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004

C:Accession: A24639; S00460; A27180; S11020; A25171; S29877; S10758

R:Shull, G.E.; Greeb, J.; Lingrel, J.B.

Biochemistry 25, 8125-8132, 1986

A:Title: Molecular cloning of three distinct forms of the Na+,K+-ATPase alpha-subunit fr

A:Accession: A24639

A:Molecule type: mRNA

A:Residues: 1-1023 &lt;SHU&gt;

A:Cross-references: UNIPROT:P06685; EMBL:M14511; NID:9203026; PIDN:AAA40775.1; PID:92030

R:Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.; R  
 J. Biochem. 102, 43-58, 1987

A:Title: Primary structures of two types of alpha-subunit of rat brain Na(+),K(+)-ATPase

A:Reference number: S00460; MUID:88032933; PMID:2822682

A:Accession: S00460

A:Molecule type: mRNA

A:Cross-references: EMBL:X05882; NID:955771; PIDN:CAA29306.1; PID:955772

R:Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.

J. Cell Biol. 105, 1855-1865, 1987

A:Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural

A:Reference number: A92749; MUID:88033255; PMID:2822726

A:Accession: A27180

A:Molecule type: mRNA

A:Residues: 1-67, 'PV', 70-174, 'E', 176-187, 'V', 189-334, 'V', 336-1023 <HER>

A:Cross-references: EMBL:M28647; NID:9205631; PIDN:AAA41671.1; PID:9205632

R:Yagawa, Y.; Kawakami, K.; Nagano, K.

Biochim. Biophys. Acta 1049, 286-292, 1990

A:Title: Cloning and analysis of the 5'-flanking region of rat Na(+)/K(+)-ATPase alpha-1

A:Reference number: S11020; MUID:90344872; PMID:2166579

A:Accession: S11020

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-41 <YAG>

A:Cross-references: EMBL:X53233

R:Schneider, J.W.; Mercer, R.W.; Caplan, M.; Emanuel, J.R.; Sweadner, K.J.; Benz Jr., B

Proc. Natl. Acad. Sci. U.S.A. 82, 6357-6361, 1985

A:Title: Molecular cloning of rat brain Na,K-ATPase alpha-subunit cDNA.

A:Reference number: A25171; MUID:85298352; PMID:2994074

A:Accession: A25171

A:Molecule type: mRNA

A:Residues: 489-533 <SCH>

R:Lytton, J.

Biochem. Biophys. Res. Commun. 132, 764-769, 1985

A:Title: The catalytic subunits of the (Na(+),K(+))-ATPase alpha and alpha(+) isozymes

A:Reference number: S29877; MUID:86050667; PMID:2998384

A:Accession: S29877

A:Status: preliminary

A:Molecule type: protein

A:Residues: 6-19 <LYT>

R:Kurihara, K.; Hosoi, K.; Kodama, A.; Ueha, T.

Biochim. Biophys. Acta 1039, 234-240, 1990

A:Title: A new electrophoretic variant of alpha subunit of Na(+)/K(+)-ATPase from the b

A:Reference number: S10758; MUID:90304196; PMID:2163680

A:Accession: S10758

A:Molecule type: protein

A:Residues: 6, 'X', 8-10, 'X', 12-16 <KUR>

A:Experimental source: submandibular gland

A>Note: designated alpha-S form; thought to arise from alpha-1 chain by post-translation

C:Genetics:

A:Gene: NKRA1

A:Introns: 4/3

C>Note: the list of introns may be incomplete

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans

F:6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status experimental <MAT>

F:6-95/Domain: intracellular #status predicted <INT1>

F:130-149/Domain: transmembrane #status predicted <TM2>

F:150-290/Domain: transmembrane #status predicted <TM2>

F:291-313/Domain: transmembrane #status predicted <TM3>

F:320-348/Domain: transmembrane #status predicted <TM4>

F:349-786/Domain: intracellular #status predicted <INT3>

F:587-783/Domain: ATPase nucleotide-binding domain homology <ATN>

F:787-810/Domain: transmembrane #status predicted <TM5>

F:849-874/Domain: transmembrane #status predicted <TM6>

F:875-952/Domain: intracellular #status predicted <INT4>

F:953-978/Domain: transmembrane #status predicted <TM7>

F:979-1023/Domain: extracellular #status predicted <EXT>

F:376/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:508/Binding site: ATP (Lys) #status predicted

F:717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1023;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 || ||: ||  
 DB 84 PTPPEWVKFC 93

## RESULT 19

S24650

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - giant toad

C/Species: Bufo marinus (giant toad)

C/Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004

C/Accession: A43451; S24650

R./Jalil, F.; Canessa, C.M.; Horisberger, J.D.; Rossier, B.C.

J. Biol. Chem. 267, 16895-16903, 1992

A./Title: Primary sequence and functional expression of a novel ouabain-resistant Na,K-ATPase

A./Reference number: A43451; MUID:92380991; PMID:1380956

A./Accession: A43451

A./Status: preliminary

A./Molecule type: mRNA

A./Residues: 1-1023 <JAI>

A./Cross-references: UNIPROT:P30714; EMBL:Z11798; NID:G62491; PIDN:CAA77842.1; PID:G62492

A./Experimental source: urinary bladder cell line TBM 18-23

A./Note: submitted to the EMBL Data Library, March 1992

A./Note: sequence extracted from NCBI backbone (NCBI:P111876)

C./Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C./Keywords: ATP, heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

F./6-1023/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>

F./6-99/Domain: intracellular #status predicted <INT1>

F./96-120/Domain: transmembrane #status predicted <TM1>

F./130-149/Domain: transmembrane #status predicted <TM2>

F./150-290/Domain: intracellular #status predicted <INT2>

F./291-313/Domain: transmembrane #status predicted <TM3>

F./320-348/Domain: transmembrane #status predicted <TM4>

F./349-786/Domain: intracellular #status predicted <INT3>

F./587-783/Domain: ATPase nucleotide-binding domain homology <ATN>

F./787-810/Domain: transmembrane #status predicted <TM5>

F./849-874/Domain: transmembrane #status predicted <TM6>

F./875-956/Domain: intracellular #status predicted <INT4>

F./953-978/Domain: transmembrane #status predicted <TM7>

F./979-1023/Domain: extracellular #status predicted <EXT>

F./376/Active site: Asp (aspartylphosphate intermediate) #status predicted

F./508/Binding site: ATP (lys) #status predicted

F./717,721,726/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1023;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 || ||: ||  
 DB 84 PTPPEWVKFC 93

## RESULT 20

A60444

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain precursor - African clawed frog

N./Alternate names: sodium pump alpha chain

C/Species: Xenopus laevis (African clawed frog)

C/Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 09-Jul-2004

C/Accession: A60444

R./Verrey, F.; Kailou, P.; Schaefer, E.; Fuentes, P.; Geering, K.; Rossier, B.C.; Kraeh

Am. J. Physiol. 256, F1034-F1043, 1989

A./Title: Primary sequence of Xenopus laevis Na(+)-K(+)-ATPase and its localization in A

A./Reference number: A60444; MUID:89285429; PMID:2544104

A./Accession: A60444

A./Status: not compared with conceptual translation

A./Molecule type: mRNA

A./Residues: 1-1025 <VER>

A./Cross-references: UNIPROT:Q92123; GB:U10108; NID:G499225; PIDN:AAA19022.1; PID:G499226

A./Note: the alpha chain is the catalytic chain.

C./Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C./Keywords: ATP, glycoprotein; hydrolase; phosphoprotein; potassium transport; sodium tr

F./589-785/Domain: ATPase nucleotide-binding domain homology <ATN>

F./217,485/Binding site: carboxylate (Asn) (covalent) #status predicted

F./376/Active site: Asp (aspartylphosphate intermediate) #status predicted

F./510/Binding site: ATP (lys) #status predicted

Query Match 50.6%; Score 43; DB 2; Length 1025;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 || ||: ||  
 DB 86 PTPPEWVKFC 95

## RESULT 21

PWCNM

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - white sucker

C/Species: Catostomus commersoni (white sucker)

C/Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C/Accession: S14740

R./Schoenrock, C.; Morley, S.D.; Okawara, Y.; Lederis, K.; Richter, D.

J. Biol. Chem. Hoppe-Seyler 372, 279-286, 1991

A./Title: Sodium and potassium ATPase of the teleost fish *Catostomus commersoni*. Sequence

A./Reference number: S14740; MUID:91282983; PMID:1711856

A./Accession: S14740

A./Molecule type: mRNA

A./Residues: 1-1027 <SCH>

A./Cross-references: UNIPROT:P25489; EMBL:X58629; NID:G62641; PIDN:CAA1483.1; PID:G62642

C./Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

C./Keywords: ATP, hydrolase, ion transport, phosphoprotein, potassium transport, sodium t

F./99-124/Domain: transmembrane #status predicted <TM1>

F./133-152/Domain: transmembrane #status predicted <TM2>

F./153-293/Domain: intracellular #status predicted <INT2>

F./294-316/Domain: transmembrane #status predicted <TM3>

F./323-351/Domain: transmembrane #status predicted <TM4>

F./352-790/Domain: intracellular #status predicted <INT3>

F./591-787/Domain: ATPase nucleotide-binding domain homology <ATN>

F./791-814/Domain: transmembrane #status predicted <TM5>

F./853-878/Domain: transmembrane #status predicted <TM6>

F./879-956/Domain: intracellular #status predicted <INT4>

F./957-987/Domain: transmembrane #status predicted <TM7>

F./983-1027/Domain: extracellular #status predicted <EXT>

F./379/Active site: Asp (aspartylphosphate intermediate) #status predicted

F./512/Binding site: ATP (lys) #status predicted

F./721,725,730/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1027;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 PTLREWISFC 14  
 || ||: ||  
 DB 87 PTPPEWVKFC 96

## RESULT 22

S03632

Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - fruit fly (*Drosophila melanogaster*)

N./Alternate names: sodium pump alpha chain

C/Species: *Drosophila melanogaster*

C/Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004

C/Accession: S03632; S07049

R./Lieberitz, R.M.; Takeya, K.; Fambrough, D.M.

EMBO J. 8, 193-202, 1989

A./Title: Molecular characterization and expression of the (Na+K)-ATPase alpha-subunit fr

A./Reference number: S03632; MUID:89231618; PMID:2540956

A./Accession: S03632

A./Molecule type: mRNA

A./Residues: 1-1038 <LEB>

A./Cross-references: UNIPROT:P13607; EMBL:X14476

A./Note: the sequence from fig. 9 is inconsistent with that from fig. 8 in having 89-Asp,



A>Note: it is uncertain whether Met-1 or Met-40 is the initiator  
 R:Varadi, A.; Gilmore-Heber, M.; Benz Jr., E.J.  
 FEBS Lett. 258, 203-207, 1989  
 A>Title: Amplification of the phosphorylation site - ATP-binding site cDNA fragment of P  
 A:Reference number: S07049; MUID:90092469; PMID:2557235  
 A:Accession: S07049  
 A:Molecule type: mRNA  
 A:Residues: 397-521 <VAR>  
 A:Cross-references: EMBL:X17471  
 A>Note: the authors translated the codon ACC for residue 3 as Asn and AAT for residue 89  
 C:Genetics:  
 A:Gene: FlyBase:Atp-alpha  
 A:Cross-references: FlyBase:Fgn0002921  
 A:Map position: 3R 93B  
 C:Superfamily: Na+/K+-transporting ATPase; alpha chain; ATPase nucleotide-binding domain  
 C:Keyword: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F:113-135/Domain: transmembrane #status predicted <TM1>  
 F:146-165/Domain: transmembrane #status predicted <TM2>  
 F:166-305/Domain: intracellular #status predicted <INT2>  
 F:306-328/Domain: transmembrane #status predicted <TM3>  
 F:335-363/Domain: transmembrane #status predicted <TM4>  
 F:364-801/Domain: intracellular #status predicted <INT3>  
 F:602-798/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:802-825/Domain: transmembrane #status predicted <TM5>  
 F:864-889/Domain: transmembrane #status predicted <TM6>  
 F:890-966/Domain: intracellular #status predicted <INT4>  
 F:967-993/Domain: transmembrane #status predicted <TM7>  
 F:994-1038/Domain: extracellular #status predicted <EXT>  
 F:391/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:523/Binding site: ATP (Lys) #status predicted  
 F:732,736,741/Active site: Asp, Asp, Lys #status predicted

Query Match 50.6%; Score 43; DB 1; Length 1038;  
 Best Local Similarity 44.4%; Pred. No. 66;  
 Matches 8; Conservative 1; Mismatches 3; Indels 6; Gaps 1;  
 QY 3 DGPTRLR-----EWISFC 14  
 DB 93 DGNLTPRKQTPPEWVKFC 110

RESULT 23  
 JH0470  
 Na+/K+-exchanging ATPase (BC 3.6.3.9) alpha chain (clone pAR1A136) - brine shrimp  
 C:Species: Artemia franciscana (brine shrimp)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
 C:Accession: JH0470; S24196  
 R:Macias, M.T.; Palmero, I.; Sastre, L.  
 Gene 105, 197-204, 1991  
 A>Title: Cloning of a cDNA encoding an Artemia franciscana Na/K ATPase alpha-subunit.  
 A:Reference number: JH0470; MUID:92039032; PMID:1657719  
 A:Accession: JH0470  
 A:Molecule type: mRNA  
 A:Residues: 1-1004 <NAC>  
 A:Cross-references: UNIPROT:P28774; EMBL:X56650; NID:g10933; PIDD:CA39972.1; PID:g10934  
 C:Superfamily: Na+/K+-transporting ATPase; alpha chain; ATPase nucleotide-binding domain  
 C:Keyword: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
 F:2-1004/Product: Na+/K+-transporting ATPase alpha chain #status predicted <NMT>  
 F:2-75/Domain: intracellular #status predicted <INT1>  
 F:76-97/Domain: transmembrane #status predicted <TM1>  
 F:111-130/Domain: transmembrane #status predicted <TM2>  
 F:131-271/Domain: intracellular #status predicted <INT2>  
 F:272-296/Domain: transmembrane #status predicted <TM3>  
 F:301-329/Domain: transmembrane #status predicted <TM4>  
 F:330-767/Domain: intracellular #status predicted <INT3>  
 F:368-764/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:768-791/Domain: transmembrane #status predicted <TM5>  
 F:830-855/Domain: transmembrane #status predicted <TM6>  
 F:856-936/Domain: intracellular #status predicted <INT4>  
 F:937-955/Domain: transmembrane #status predicted <TM7>  
 F:956-1004/Domain: extracellular #status predicted <EXT>  
 F:357/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:489/Binding site: ATP (Lys) #status predicted

F:698,702,707/Active site: Asp, Asp, Lys #status predicted

Query Match 50.0%; Score 42.5; DB 2; Length 1004;  
 Best Local Similarity 47.4%; Pred. No. 77;  
 Matches 9; Conservative 0; Mismatches 3; Indels 7; Gaps 1;  
 QY 3 DGP-----TLREWISFC 14  
 DB 55 DGNCLTPRKTPPEWVKFC 73

RESULT 24  
 F6876  
 hypothetical protein yuJa [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C:Species: Lactococcus lactis subsp. lactis  
 C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: F6876  
 R:Bolotin, A.; Winkler, P.; Mueger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrl  
 Genome Res. 11, 731-753, 2001  
 A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
 A:Reference number: A86625; MUID:21235186; PMID:11374771  
 A:Accession: F6876  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-312 <STO>  
 A:Cross-references: UNIPROT:Q9CE34; GB:AE005176; PID:g12725061; PIDD:AAK06112.1; GSPDB:  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: yuJa

Query Match 49.4%; Score 42; DB 2; Length 312;  
 Best Local Similarity 77.8%; Pred. No. 30;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GPTLRBWS 12  
 DB 242 GPLKKEWS 250

RESULT 25  
 D69226  
 hypothetical protein MTH943 - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Dec-2002  
 C:Accession: D69226  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N  
 Kl. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: D69226  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-522 <MTH>  
 A:Cross-references: GB:AE000868; GB:AE000666; NID:g2622025; PIDD:AB85441.1; PID:g26220  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH943  
 A:Start codon: GTG  
 C:Superfamily: uncharacterized conserved protein

Query Match 49.4%; Score 42; DB 2; Length 522;  
 Best Local Similarity 70.0%; Pred. No. 50;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 ADGPTLRW 11  
 DB 198 ADGTVYEW 207

RESULT 26  
 S62941

probable membrane protein YNL029c - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: hypothetical protein N2755  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004  
 C:Accession: S62941; S62951  
 R:André, B.; Iraqi Housseini, I.; Urrestazu, L.A.; Vissers, S.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S62920  
 A:Accession: S62941  
 A:Molecule type: DNA  
 A:Residues: 1-522 <AND>  
 A:Cross-references: UNIPROT:P53966; EMBL:Z71305; NID:g1301864; PID:g1301865; MIPS:YNL029  
 A:Experimental source: strain S288C  
 R:Duisterhoft, A.; Floeth, M.; Fritze, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S62944  
 A:Accession: S62951  
 A:Molecule type: DNA  
 A:Residues: 1-522 <DUB>  
 A:Cross-references: EMBL:Z71305; NID:g1301864; PID:g1301865; MIPS:YNL029C  
 A:Experimental source: strain S288C  
 C:Keywords: transmembrane protein  
 A:Map position: 14L  
 C:Keywords: transmembrane #status predicted <TMM>  
 F:21-37/Domain

Query Match 49.4%; Score 42; DB 2; Length 522;  
 Best Local Similarity 50.0%; Pred. No. 50;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 ADGPTLRWISF 13  
 : ||| : ||| :  
 DB 285 SDDPELRDWINY 236

RESULT 27  
 A1544  
 conserved hypothetical protein lin0897 [imported] - *Listeria innocua* (strain Clijp11262)  
 C:Species: *Listeria innocua*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: A11544  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Ertian, K.D.; Fshih, H. D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma Ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomes of *Listeria species*.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: A11544  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-725 <GLA>  
 A:Cross-references: UNIPROT:Q92D58; GB:ML52022; PIDN:CA06129.1; PID:g16413347; GSPDB:G  
 A:Experimental source: strain Clijp11262  
 C:Genetics:  
 A:Gene: lin0897  
 C:Superfamily: hypothetical protein ydci

Query Match 49.4%; Score 42; DB 2; Length 725;  
 Best Local Similarity 70.0%; Pred. No. 68;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ADGPTLRWISF 11  
 : ||| : ||| :  
 DB 174 SDEPTLRWISF 183

RESULT 28  
 T12091  
 starch phosphorylase (EC 2.4.1.1) H, cytosolic isoform - fava bean

N:Alternate names: alpha 1,4-glucan phosphorylase type H  
 C:Species: *Vicia faba* (fava bean)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:Accession: T12091  
 R:Buchner, P.; Borisjuk, L.; Wobus, U.  
 Planta 199, 64-73, 1996  
 A:Title: Glucan phosphorylases in *Vicia faba* L.: cloning, structural analysis and expres  
 A:Reference number: Z17412; MUID:96236831; PMID:8680306  
 A:Accession: T12091  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-842 <BUC>  
 A:Cross-references: UNIPROT:P53337; EMBL:Z35117; NID:9510931; PIDN:CA84494.1; PID:95109  
 A:Experimental source: strain *Vicia faba* var. minor; cultivar Fribo; coveledon, clone VF  
 C:Genetics:  
 A:Gene: Pho2  
 C:Function:  
 A:Description: catalyzes the formation of glucose 1-phosphate from polyglucose  
 C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphate  
 C:Superfamily: glucan phosphorylase  
 A:Reference number: F688; Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 49.4%; Score 42; DB 2; Length 842;  
 Best Local Similarity 58.3%; Pred. No. 78;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DGPTLRWISFC 14  
 : ||| : ||| :  
 DB 488 NGITPRRWINFC 499

RESULT 29  
 S07755  
 hypothetical protein 16 - *Paramecium tetraurelia* mitochondrion  
 C:Species: *mitochondrion Paramecium tetraurelia*  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004  
 C:Accession: S07755  
 R:Pritchard, A.E.; Sellhammer, J.J.; Mahalingam, R.; Sable, C.L.; Vennu, S.E.; Cummings, Nucleic Acids Res. 18, 173-180, 1990  
 A:Title: Nucleotide sequence of the mitochondrial genome of *Paramecium*.  
 A:Reference number: S07725; MUID:90174913; PMID:2308823  
 A:Accession: S07755  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-189 <PRI>  
 A:Cross-references: UNIPROT:P15617; EMBL:X15917; NID:g13256; PID:g578768  
 C:Genetics:  
 A:Gene: mitochondrial  
 A:Genetic code: SGC6  
 A:Start codon: ATT  
 C:Keywords: mitochondrion

Query Match 48.2%; Score 41; DB 2; Length 189;  
 Best Local Similarity 72.7%; Pred. No. 27;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 DGPTLRWISF 13  
 : ||| : ||| :  
 DB 92 DEPTLRWISF 102

RESULT 30  
 H70849  
 hypothetical protein RV0079 - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: H70849  
 R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. i; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome



A:Reference number: A70500; PMID:98295987; PMID:9634230  
 A:Accession: H70849  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-273 <COL>  
 A:Cross-references: UNIPROT:Q53624; GB:AL021428; GB:AL123456; NID:g3261514; P1DN:CA1626  
 A:Experimental source: strain H378v  
 C:Genetics:  
 A:Gene: RV0079  
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0079

Query Match 48.2%; Score 41; DB 2; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CADPTL 7  
 |||||  
 Db 66 CADPTL 72

RESULT 31  
 E84853  
 hypothetical protein At2g42400 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: E84853  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bent, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 Neus, D.; Newman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; PMID:20083487; PMID:10617197  
 A:Accession: E84853  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1473 <SNO>  
 A:Cross-references: UNIPROT:Q9SLB9; GB:AE002093; NID:g4567312; P1DN:AD23723.1; GSPDB:GN  
 C:Genetic:  
 A:Gene: At2g42400  
 A:Map position: 2

Query Match 48.2%; Score 41; DB 2; Length 473;  
 Best Local Similarity 54.5%; Pred. No. 66;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DGPTLRWISF 13  
 :|||:  
 Db 344 EGETIRWLPF 354

RESULT 32  
 T10947  
 starch phosphorylase (EC 2.4.1.1) precursor - sweet potato  
 C:Species: Ipomoea batatas (sweet potato)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:Accession: T10947  
 R:Lin, C.T.; Yeh, K.W.; Lee, P.D.; Su, J.C.  
 submitted to the EMBL Data Library, July 1991  
 A>Description: Primary structure of sweet potato starch phosphorylase deduced from its c  
 A:Reference number: Z17224  
 A:Accession: T10947  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-955 <LIN>  
 A:Cross-references: UNIPROT:P27598; EMBL:M64362; NID:g168275; PID:g168276  
 A:Experimental source: cv. Tainong 57; tuberous root  
 C:Genetics:  
 A:Genome: nuclear  
 A>Note: starch phosphorylase  
 C:Function:  
 A>Description: catalyzes the formation of glucose 1-phosphate from polyglucose  
 C:Superfamily: glucan phosphorylase  
 C:Keywords: chloroplast; glycosyltransferase; hexosyltransferase; phosphoprotein; pyridic

F/1-43/Domain: transit peptide (chloroplast) #status predicted <NP>  
 F/44-95/Product: starch phosphorylase #status predicted <MAT>  
 F/801/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 48.2%; Score 41; DB 2; Length 955;  
 Best Local Similarity 58.3%; Pred. No. 13e+02;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DGPTLRWISFC 14  
 :|||:  
 Db 600 NGVTPRRWIRFC 611

RESULT 33  
 PHPOAG  
 starch phosphorylase (EC 2.4.1.1) precursor - potato  
 N:Alternate names: alpha-glucan phosphorylase  
 C:Species: Solanum tuberosum (potato)  
 C>Date: 04-Dec-1986 #sequence\_revision 30-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: U00130; A00574; F00139; S15531; S12033  
 R:Nakano, K.; Mori, H.; Fukui, T.  
 J. Biochem. 106, 691-695, 1989  
 A>Title: Molecular cloning of cDNA encoding potato amyloplast alpha-glucan phosphorylase.  
 A:Reference number: A91915; PMID:90110071; PMID:2481677  
 A:Accession: U00130  
 A:Molecule type: mRNA  
 A:Residues: 1-966 <NAL>  
 A:Cross-references: UNIPROT:P04045; GB:D00520; NID:g3702676; P1DN:BA00407.1; PID:g2179  
 R:Nakano, K.; Fukui, T.  
 J. Biol. Chem. 261, 8230-8236, 1986  
 A>Title: The complete amino acid sequence of potato alpha-glucan phosphorylase.  
 A:Reference number: A92591; PMID:86250715; PMID:3722153  
 A:Accession: A00574  
 A:Molecule type: protein  
 A:Residues: 51-966 <NA>  
 R:Brison, N.; Groux, H.; Zollinger, M.; Camirand, A.; Simard, C.  
 Plant Cell 1, 559-566, 1989  
 A>Title: Maturation and subcellular compartmentation of potato starch phosphorylase.  
 A:Reference number: P00139; PMID:92404721; PMID:2535551  
 A:Accession: P00139  
 A:Molecule type: mRNA  
 A:Residues: 1-130 <BRI>  
 A:Experimental source: tuber, cv. Kennebec  
 R:Brison, N.  
 submitted to the EMBL Data Library, April 1990  
 A:Reference number: S15531  
 A:Accession: S15531  
 A:Molecule type: mRNA  
 A:Residues: 1158 'D', 160-966 <BR2>

A:Cross-references: EMBL:X52385; NID:g21578; P1DN:CA36612.1; PID:g21579  
 R:Camirand, A.; St-Pierre, B.; Marneau, C.; Brison, N.  
 Mol. Gen. Genet. 224, 33-39, 1990  
 A>Title: Occurrence of a copia-like transposable element in one of the introns of the p  
 A:Reference number: S12033; PMID:91117174; PMID:1703627  
 A:Accession: S12033  
 A:Molecule type: mRNA  
 A:Residues: 416-595 <CAN>  
 A:Cross-references: EMBL:X52385  
 C:Comment: Phosphorylase, an important allosteric enzyme in carbohydrate metabolism, cat  
 gulatory mechanisms and in their natural substrates. However, all known phosphorylaes  
 C:Superfamily: glucan phosphorylase  
 C:Keywords: allosteric regulation; carbohydrate metabolism; glycosyltransferase; hexose  
 F/1-50/Domain: transit peptide (amyloplast) #status predicted <MAT>  
 F/51-966/Product: phosphorylase #status experimental <MAT>  
 F/812/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 48.2%; Score 41; DB 1; Length 966;  
 Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DGPTLRWISFC 14  
 :|||:  
 Db 611 NGVTPRRWIRFC 622

RESULT 34  
T09210  
starch phosphorylase (EC 2.4.1.1) - spinach  
M:Alternate names: alpha-glucan phosphorylase  
C:Species: Spinacia oleracea (spinach)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 18-Aug-2003  
C:Accession: T09210  
R:Duvenig, E.; Steud, M.; Wilmitzer, L.; Kossmann, J.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: Z16608  
A:Accession: T09210  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-971 <DUM>  
A:Cross-references: EMBL:X85181  
A:Experimental source: U.S. hybrid 424 Serry-Mores seed company (CA, USA); flower  
C:Genetics:  
A:Gene: SPP11  
C:Function:  
A:Description: catalyzes the formation of glucose 1-phosphate from polyglucose  
C:Superfamily: glucan phosphorylase  
C:Keywords: carbohydrate metabolism; glycosyltransferase; hexosyltransferase; phosphop  
F:817/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 48.2%; Score 41; DB 2; Length 971;  
Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGPTRLREWISFC 14  
DB 616 NGVTPRRWIRFC 627

RESULT 35  
S47243  
starch phosphorylase (EC 2.4.1.1) isoform L precursor, chloroplast - fava bean  
M:Alternate names: alpha-1,4 glucan phosphorylase isoform L  
C:Species: Vicia faba (fava bean)  
C:Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 18-Aug-2003  
C:Accession: S47243  
R:Buchner, P.; Bortisjuk, L.; Wobus, U.  
submitted to the EMBL Data Library, August 1994  
A:Description: alpha-1,4 glucan phosphorylases in Vicia faba L.: cDNA-Characterization a  
A:Reference number: S47243  
A:Accession: S47243  
A:Molecule type: mRNA  
A:Residues: 1-1000 <BUC>  
A:Cross-references: EMBL:Z36880  
A:Experimental source: strain Vicia faba var. minor; cultivar Fribo; cotyledon; clone VF  
C:Genetics:  
A:Gene: Phol  
A:Genome: nuclear  
C:Superfamily: glucan phosphorylase  
C:Keywords: chloroplast; glycosyltransferase; hexosyltransferase; phosphoprotein; pyrid  
F:1-61/Domain: transit peptide (chloroplast) #status predicted <TMP>  
F:62-1000/Product: starch phosphorylase isoform L #status predicted <MAT>  
F:846/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 48.2%; Score 41; DB 2; Length 1000;  
Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGPTRLREWISFC 14  
DB 645 NGVTPRRWIRFC 656

RESULT 36  
T17884  
S-layer protein - Bacillus circulans  
C:Species: Bacillus circulans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T17884  
R:Aubert-Pivert, E.; Davies, J.  
Gene 147, 1-11, 1994  
A:Title: Biosynthesis of butirosin in Bacillus circulans NRRL B3112: Identification by  
A:Reference number: Z18808; MUID:94374689; PMID:7522196  
A:Accession: T17884  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1616 <AUB>  
A:Cross-references: UNIPROT:P35824; EMBL:L20421; NID:g304142; PID:g304143; PIDN:AAA6258  
C:Genetics:  
A:Gene: butB  
C:Function:  
A:Pathway: butirosin biosynthesis

Query Match 48.2%; Score 41; DB 2; Length 1616;  
Best Local Similarity 54.5%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DGPTRLREWISFC 13  
DB 737 DGPTRLRWMEF 747

RESULT 37  
A70301  
ribosomal protein L22 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 13-Aug-1999  
C:Accession: A70301  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: A70301  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-98 <AQF>  
A:Cross-references: GB:AE000669; NID:g2982762; PIDN:AA06396.1; PID:g2982770; GB:AE0006  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: rplV  
C:Superfamily: Escherichia coli ribosomal protein L22

Query Match 47.1%; Score 40; DB 2; Length 98;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 DGPTRLREWT 11  
DB 59 DGPTRLKWI 67

RESULT 38  
S21826  
T-cell receptor beta chain V region homolog - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S21826  
R:George Jr., J.F.; Schroeder Jr., H.W.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S21826  
A:Accession: S21826  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-152 <GEO>  
A:Cross-references: EMBL:X56142; NID:g37500; PIDN:CAA39607.1; PID:g388518  
C:Genetics:  
A:Insertions: 134/3; 139/2  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:56-133/Domain: immunoglobulin homology <IMH>

Query Match 47.1%; Score 40; DB 2; Length 152;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 CAGDPTLRWISFC 14  
 DB 20 CAGPGILLCWVLLC 33

## RESULT 39

ICMS2

Interleukin-2 precursor - mouse

N/Alternate names: IL-2; T-cell growth factor (TCGF)

C/Species: Mus musculus (house mouse)

C/Date: 30-Jun-1987 #sequence, revision 30-Jun-1987 #text, change 09-Jul-2004

C/Accession: A93550; A54490; A94064; I48597; A01850; I84713

R/Phase: A; Fujita, T.; Yasumitsu, H.; Kashima, N.; Hasegawa, K.; Taniguchi, T.

Nucleic Acids Res. 12, 9323-9331, 1984

A/Title: Organization and structure of the mouse interleukin-2 gene.

A/Reference number: A93550; PMID:85087940; PMID:6240025

A/Accession: A93550

A/Molecule type: DNA

A/Residues: 1-169 <RUS>

A/Cross-references: UNIPROT:P04351

R/Degrave, W.; Simons, G.; Devos, R.; Plaetinck, G.; Remaut, E.; Tavernier, J.; Fiers, W.

Mol. Biol. Rep. 11, 57-61, 1986

A/Title: Cloning and structure of a mouse interleukin-2 chromosomal gene.

A/Reference number: A54490; PMID:86118396; PMID:3003564

A/Accession: A54490

A/Molecule type: DNA

A/Residues: 1-169 <DEG>

A/Cross-references: GB:M16760

R/Yokota, T.; Arai, N.; Lee, F.; Rennick, D.; Mosmann, T.; Arai, K.

Proc. Natl. Acad. Sci. U.S.A. 82, 68-72, 1985

A/Title: Use of a cDNA expression vector for isolation of mouse interleukin 2 cDNA clone

A/Reference number: A94064; PMID:85113172; PMID:3918306

A/Accession: A94064

A/Molecule type: mRNA

A/Residues: 1-169 <YOK>

A/Cross-references: GB:K02292; NID:G198330; PIDN:AAA9289.1; PID:9309404

R/Kashima, N.; Nishit-Takaoka, C.; Fujita, T.; Taki, S.; Yamada, G.; Hamuro, J.; Taniguchi

Nature 313, 402-404, 1985

A/Title: Unique structure of murine interleukin-2 as deduced from cloned cDNAs.

A/Reference number: I48597; PMID:85111148; PMID:2578624

A/Accession: I48597

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-169 <RES>

A/Cross-references: EMBL:X01772; GB:K02797; NID:G52663; PIDN:CAA25909.1; PID:9758159

C/Comment: Produced by T-cells in response to antigenic or mitogenic stimulation, this p

C/Genetics:

A/Introns: 63/3; 83/3; 132/3

C/Superfamily: interleukin-2

C/Keywords: cytokine; glycoprotein; growth factor; immunoregulation; lymphokine; T-cell

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-169/Product: interleukin-2 #status predicted <MAT>

F/23/Binding site: carbohydrate (Thr) (covalent) #status predicted

F/92-140/Disulfide bond: #status predicted

C/Date: 13-Jan-1995 #sequence, revision 13-Jan-1995 #text, change 09-Jul-2004

C/Accession: S37289; S27205; S36162; S24936

R/Jodd, U.A.

submitted to the EMBL Data Library, April 1993

A/Reference number: S37289

A/Accession: S37289

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-169 <TOD>

A/Cross-references: UNIPROT:Q8BHA4; EMBL:X73040

R/Matesanz, F.; Alcina, A.; Pellicer, A.

Biochim. Biophys. Acta 1132, 335-336, 1992

A/Title: A new cDNA sequence for the murine interleukin-2 gene.

A/Reference number: S27205; PMID:93041941; PMID:1420317

A/Accession: S27205

A/Molecule type: mRNA

A/Residues: 1-63 <MTE>

A/Cross-references: EMBL:X66058; NID:G52725; PIDN:CAA46854.1; PID:952726

R/Ghosh, S.; Palmer, S.M.; Rodrigues, N.R.; Cordell, H.J.; Hearne, C.M.; Cornall, R.J.

Nature Genet. 4, 404-409, 1993

A/Title: Polygenic control of autoimmune diabetes in nonobese diabetic mice.

A/Reference number: S36162; PMID:94004970; PMID:8401590

A/Accession: S36162

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-50 <GHO>

A/Cross-references: EMBL:X73040

C/Superfamily: interleukin-2

C/Keywords: cytokine; glycoprotein; growth factor; lymphokine; T-cell

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-63/Product: interleukin-2 #status predicted <MAT>

Query Match 47.1%; Score 40; DB 2; Length 169;  
 Best Local Similarity 75.0%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 LREMISFC 14  
 DB 153 LRRMIAPC 160

## RESULT 41

E95908

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSym

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence, revision 24-Aug-2001 #text, change 09-Jul-2004

C/Accession: E95908

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing end

A/Reference number: A95842; PMID:21396508; PMID:11481431

A/Accession: E95908

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-169 <KIR>

A/Cross-references: UNIPROT:Q92W14; GB:AL591985; PIDN:CAC48933.1; PID:915140418; GSPDB:

A/Experimental source: strain 1021, megaplasmid pSymB

R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Aboja, P.; Ampe, F.; Barloy-Hubler

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure

hebnick, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; PMID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Gene: SMO20554

A/Genome: plasmid

C/Superfamily: uncharacterized conserved protein

Query Match 47.1%; Score 40; DB 2; Length 169;  
 Best Local Similarity 53.8%; Pred. No. 36;

Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CADPTLRWISF 13  
| | | | |  
| | | | |  
Db 52 CODDPTRIRRCIGF 64

RESULT 42  
S46354  
pol polyprotein - simian immunodeficiency virus sivagm (isolate SABD37) (fragment)  
C:Species: simian immunodeficiency virus sivagm  
A:Variety: isolate SABD37  
C:Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 26-Aug-1999  
C:Accession: S46354  
R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;  
EMBO J. 13, 2935-2947, 1994  
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African gree  
A:Reference number: S46335; MUID:94298785; PMID:8026477  
A:Accession: S46354  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-217 <1IN>  
A:Cross-references: EMBL:U04018; NID:g466250; PIDN:AAA21512.1; PID:g466251  
A:Experimental source: isolate SABD37, sabaus monkey  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993  
C:Genetics:  
A:Gene: pol  
C:Superfamily: pol polyprotein  
C:Keywords: polyprotein

Query Match 47.1%; Score 40; DB 2; Length 217;  
Best Local Similarity 75.0%; Pred. No. 45;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DGPTLRW 10  
| | | | |  
| | | | |  
Db 86 DGPTLRW 93

RESULT 43  
B97072  
probable hydrolase of PHP superfamily [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: B97072  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: B97072  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-252 <RUR>  
A:Cross-references: UNIPROT:Q97J89; GB:AE001437; PIDN:AAK79365.1; PID:g15024335; GSPDB:C  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1397  
C:Superfamily: Mg-dependent DNase, Tsd type

Query Match 47.1%; Score 40; DB 2; Length 252;  
Best Local Similarity 60.0%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 DGPTLRWIS 12  
| | | | |  
| | | | |  
Db 197 DGPTLRWIS 206

RESULT 44  
B69096  
corrinoid/iron-sulfur protein, small subunit - Methanobacterium thermoautotrophicum (str  
C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: B69096  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
J. Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
Kl, S.; Church, G.M.; Daniele, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514; PMID:93771463  
A:Accession: B69096  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-389 <MTH>  
A:Cross-references: UNIPROT:Q27747; GB:AE000928; GB:AE000666; NID:g2622835; PIDN:AA8611  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1712  
A:Start codon: GTG  
C:Superfamily: corrinoid/iron-sulfur protein small chain

Query Match 47.1%; Score 40; DB 2; Length 389;  
Best Local Similarity 63.6%; Pred. No. 79;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ADGPTLRWIS 12  
| | | | |  
| | | | |  
Db 372 ADGPTLRWIS 382

RESULT 45  
S36113  
LIS-1 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 16-Aug-2004  
C:Accession: S36113  
R:Reiner, O.; Carozzo, R.; Shen, Y.; Wehnert, M.; Faustiniella, F.; Dobyns, W.B.; Caskey  
Nature 364, 717-721, 1993  
A:Title: Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-subu  
A:Reference number: S36113; MUID:93361119; PMID:8355785  
A:Accession: S36113  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-409 <RAS>  
C:Genetics:  
A:Gene: LIS-1  
C:Superfamily: WD repeat homology  
F:103-136/Domain: WD repeat homology <WD1>  
F:145-178/Domain: WD repeat homology <WD2>  
F:187-220/Domain: WD repeat homology <WD3>  
F:229-262/Domain: WD repeat homology <WD4>  
F:333-366/Domain: WD repeat homology <WD5>  
F:375-408/Domain: WD repeat homology <WD6>

Query Match 47.1%; Score 40; DB 2; Length 409;  
Best Local Similarity 70.0%; Pred. No. 83;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CADGPTLRW 10  
| | | | |  
| | | | |  
Db 355 CADGPTLRW 364

Search completed: September 1, 2005, 16:23:00  
Job time : 11.6763 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 52.0719 Seconds  
(without alignments)  
137.677 Million cell updates/sec

Title: US-10-083-768-12

Perfect score: 85  
Sequence: 1 CADGPTLRWISFC 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	62.4	297	2 Q7U0E4	Q7U0E4 rhodopirell
2	50.5	59.4	387	2 Q98A97	Q98A97 rhizobium 1
3	50.5	59.4	389	2 Q8KJF9	Q8KJF9 rhizobium 1
4	48	56.5	1123	2 Q7OC63	Q7OC63 anophelis g
5	47.5	55.9	283	2 Q82CW2	Q82CW2 streptomyc
6	46	54.1	319	2 Q9RKM5	Q9RKM5 streptomyc
7	46	54.1	347	2 Q7PPE6	Q7PPE6 anophelis g
8	45	52.9	108	2 Q7RUA5	Q7RUA5 neurospora
9	45	52.9	173	2 Q8C4M6	Q8C4M6 mus musculu
10	45	52.9	209	2 Q6NIX5	Q6NIX5 rhodopseudo
11	45	52.9	309	2 Q8XZM5	Q8XZM5 ralsstonia s
12	45	52.9	443	2 Q9P8S8	Q9P8S8 phaeophaer
13	44	51.8	173	2 Q6QHD2	Q6QHD2 galliid herp
14	44	51.8	178	2 Q6PLI4	Q6PLI4 galliid herp
15	44	51.8	209	2 Q9LO59	Q9LO59 streptomyc
16	44	51.8	292	2 Q67642	Q67642 galliid herp
17	44	51.8	298	2 Q86653	Q86653 galliid herp
18	44	51.8	974	1 PHS2_SOLTU	P53535 solanum tub
19	44	51.8	997	2 Q6B1Z6	Q6B1Z6 debaryomyce
20	44	51.8	1008	2 Q8AY57	Q8AY57 fundulus he
21	44	51.8	1011	2 Q6VYM7	Q6VYM7 oncorhynch
22	44	51.8	1022	1 ATIA_TORCA	P05025 torpedo cal
23	44	51.8	1023	1 A1A1_HUMAN	P05023 homo sapien
24	44	51.8	1025	2 Q7ZYF8	Q7ZYF8 xenopus lae
25	44	51.8	1028	1 A1A4_RAT	Q64541 rattus norv
26	43.5	51.2	405	2 Q9KIE9	Q9KIE9 streptomyc
27	43.5	51.2	934	2 Q9NEX6	Q9NEX6 caenorhabdi
28	43	50.6	127	2 Q9N0Z5	Q9N0Z5 oryctolagus
29	43	50.6	171	2 Q8HYW6	Q8HYW6 bos taurus
30	43	50.6	176	2 Q866A9	Q866A9 equus cabal
31	43	50.6	245	2 Q9M060	Q9M060 arabidopsis

32	43	50.6	407	2 Q6NMU4	Q6NMU4 drosophila
33	43	50.6	407	2 Q9VK55	Q9VK55 drosophila
34	43	50.6	469	2 Q37839	Q37839 bacterioph
35	43	50.6	490	2 Q04270	Q04270 chlamydom
36	43	50.6	509	2 Q8B1G9	Q8B1G9 mus musculu
37	43	50.6	960	2 Q80U28	Q80U28 mus musculu
38	43	50.6	962	2 Q91YY9	Q91YY9 mus musculu
39	43	50.6	1000	2 Q7Z4I9	Q7Z4I9 homo sapien
40	43	50.6	1009	2 Q96SL3	Q96SL3 electrophor
41	43	50.6	1010	1 A1A3_CHICK	P24798 gallus gall
42	43	50.6	1010	1 A1A3_OREMO	P58312 oreochromis
43	43	50.6	1012	2 Q6VYM8	Q6VYM8 oncorhynch
44	43	50.6	1013	1 A1A3_HUMAN	P13637 homo sapien
45	43	50.6	1013	1 A1A3_MOUSE	Q6P1C6 mus musculu
46	43	50.6	1013	1 A1A3_RAT	P06687 rattus norv
47	43	50.6	1017	1 A1A2_CHICK	P24797 gallus gall
48	43	50.6	1017	2 Q9DX34	Q9DX34 brachydanio
49	43	50.6	1017	2 Q9DGL5	Q9DGL5 brachydanio
50	43	50.6	1020	1 A1A2_HUMAN	P50993 homo sapien
51	43	50.6	1020	1 A1A2_RAT	P06686 rattus norv
52	43	50.6	1020	2 Q6P1E5	Q6P1E5 mus musculu
53	43	50.6	1020	2 Q6PAG0	Q6PAG0 xenopus lae
54	43	50.6	1021	1 A1A1_CANFA	P50997 canis famill
55	43	50.6	1021	1 A1A1_CHICK	P09572 gallus gall
56	43	50.6	1021	1 A1A1_HORSE	P18907 equus cabal
57	43	50.6	1021	1 A1A1_PIG	P05024 sus scrofa
58	43	50.6	1021	1 A1A1_SHEEP	P04074 ovis aries
59	43	50.6	1022	1 A1A1_ANGAN	Q92010 anguilla an
60	43	50.6	1022	2 Q6ZQ49	Q6ZQ49 mus musculu
61	43	50.6	1022	2 Q90WE7	Q90WE7 carassius a
62	43	50.6	1023	1 A1A1_BUPMA	P10714 bufo martinu
63	43	50.6	1023	1 A1A1_MOUSE	Q8VDJ2 mus musculu
64	43	50.6	1023	1 A1A1_OREMO	Q9YH26 oreochromis
65	43	50.6	1023	1 A1A1_RAT	P06685 rattus norv
66	43	50.6	1023	2 Q9N0Z6	Q9N0Z6 oryctolagus
67	43	50.6	1023	2 Q6P1S7	Q6P1S7 xenopus tro
68	43	50.6	1023	2 Q6P271	Q6P271 brachydanio
69	43	50.6	1023	2 Q7ZSX5	Q7ZSX5 xenopus lae
70	43	50.6	1023	2 Q7ZYV1	Q7ZYV1 anas platyr
71	43	50.6	1023	2 Q8AY58	Q8AY58 fundulus he
72	43	50.6	1023	2 Q9DEU2	Q9DEU2 brachydanio
73	43	50.6	1023	2 Q9DEY2	Q9DEY2 brachydanio
74	43	50.6	1024	2 Q90X33	Q90X33 brachydanio
75	43	50.6	1024	2 Q9DEU0	Q9DEU0 xenopus lae
76	43	50.6	1025	1 A1A1_XENTLA	Q92123 xenopus lae
77	43	50.6	1025	2 Q6IP41	Q6IP41 xenopus lae
78	43	50.6	1025	2 Q6VTM6	Q6VTM6 oncorhynch
79	43	50.6	1025	2 Q9DEU1	Q9DEU1 brachydanio
80	43	50.6	1027	1 A1A1_CANTCO	P25489 catostomus
81	43	50.6	1028	2 Q6VTM5	Q6VTM5 oncorhynch
82	43	50.6	1028	2 Q7ZU25	Q7ZU25 brachydanio
83	43	50.6	1028	2 Q9DGL6	Q9DGL6 brachydanio
84	43	50.6	1029	1 A1A4_HUMAN	Q13733 homo sapien
85	43	50.6	1032	1 A1A4_MOUSE	Q9WV27 mus musculu
86	43	50.6	1053	2 Q8VCE0	Q8VCE0 mus musculu
87	43	50.6	2098	2 Q7S687	Q7S687 neurospora
88	42.5	50.0	322	2 Q9U514	Q9U514 attemia par
89	42.5	50.0	394	2 Q9U516	Q9U516 attemia san
90	42.5	50.0	399	2 Q9U605	Q9U605 attemia san
91	42.5	50.0	454	2 Q9U606	Q9U606 attemia san
92	42.5	50.0	454	2 Q9U607	Q9U607 attemia san
93	42.5	50.0	454	2 Q9U608	Q9U608 attemia san
94	42.5	50.0	454	2 Q9U609	Q9U609 attemia san
95	42.5	50.0	454	2 Q9U610	Q9U610 attemia san
96	42.5	50.0	454	2 Q9U611	Q9U611 attemia san
97	42.5	50.0	454	2 Q74240	Q74240 chleavla h
98	42.5	50.0	1004	1 A1B1_ARSPF	P28774 attemia san
99	42.5	50.0	1004	2 Q9V492	Q9V492 drosophila
100	42	49.4	168	2 Q9V492	Q9V492 drosophila

## ALIGNMENTS

Q6NMU4	drosophila
Q9VK55	drosophila
Q37839	bacterioph
Q04270	chlamydom
Q8B1G9	mus musculu
Q80U28	mus musculu
Q91YY9	mus musculu
Q7Z4I9	homo sapien
Q96SL3	electrophor
P24798	gallus gall
P58312	oreochromis
Q6VYM8	oncorhynch
P13637	homo sapien
Q6P1C6	mus musculu
P06687	rattus norv
P24797	gallus gall
Q9DX34	brachydanio
Q9DGL5	brachydanio
P50993	homo sapien
P06686	rattus norv
Q6P1E5	mus musculu
Q6PAG0	xenopus lae
P50997	canis famill
P09572	gallus gall
P18907	equus cabal
P05024	sus scrofa
P04074	ovis aries
Q92010	anguilla an
Q6ZQ49	mus musculu
Q90WE7	carassius a
P10714	bufo martinu
Q8VDJ2	mus musculu
Q9YH26	oreochromis
P06685	rattus norv
Q9N0Z6	oryctolagus
Q6P1S7	xenopus tro
Q6P271	brachydanio
Q7ZSX5	xenopus lae
Q7ZYV1	anas platyr
Q8AY58	fundulus he
Q9DEU2	brachydanio
Q9DEY2	brachydanio
Q90X33	brachydanio
Q9DEU0	xenopus lae
Q92123	xenopus lae
Q6IP41	xenopus lae
Q6VTM6	oncorhynch
Q9DEU1	brachydanio
P25489	catostomus
Q6VTM5	oncorhynch
Q7ZU25	brachydanio
Q9DGL6	brachydanio
Q13733	homo sapien
Q9WV27	mus musculu
Q8VCE0	mus musculu
Q7S687	neurospora
Q9U514	attemia par
Q9U516	attemia san
Q9U605	attemia san
Q9U606	attemia san
Q9U607	attemia san
Q9U608	attemia san
Q9U609	attemia san
Q9U610	attemia san
Q9U611	attemia san
Q74240	chleavla h
P28774	attemia san
Q9V492	drosophila

```

RESULT 1
Q7UOE4 PRELIMINARY; PRT; 297 AA.
ID Q7UOE4;
AC Q7UOE4;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=RB6375;
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Hellmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294144; CAD74759.1; -.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00829; GYP_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475670F02C78E9B CRC64;

Query Match
Best Local Similarity 62.4%; Score 53; DB 2; Length 297;
Matches 8; Conservative 72.7%; Pred. No. 0.94;
Mismatch 0; Indels 0; Gaps 0;

QY 2 ADGPTLRWIS 12
DB 175 ADGPTLRWIS 185

RESULT 2
Q98A97 PRELIMINARY; PRT; 387 AA.
ID Q98A97;
AC Q98A97;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE M1r6096 protein.
GN OrderedlocusNames=mlr6096;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matsumoto M., Itoh Y., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Wochizuki Y., Nakayama S., Kohazaki N., Shimizu S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003008; BAB52440.1; -.
DR HSP; P77407; IPOY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome.
SQ SEQUENCE 387 AA; 42226 MW; 64643EBEC8F25518 CRC64;

```

```

Query Match
Best Local Similarity 59.4%; Score 50.5; DB 2; Length 387;
Matches 9; Conservative 42.9%; Pred. No. 3.4;
Mismatch 2; Indels 7; Gaps 1;

QY 1 CADGPTL-----REWISFC 14
DB 243 CADGKEVIFSVQNDREWNFC 257

RESULT 3
Q8KJF9 PRELIMINARY; PRT; 389 AA.
ID Q8KJF9;
AC Q8KJF9;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE PUTATIVE RACEMASE/DEHYDRATASE PROTEIN.
GN Name=msl181;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R7A;
RX MEDLINE=2199272; PubMed=12003951;
RX DOI=10.1128/JB.184.11.3086-3095.2002;
RA Sullivan J.T., Trebatowski J.R., Crickshank R.W., Gouzy J.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rosbach U.,
RA Stuart G.S., Weaver J.E., Webb R.J., de Brito F.J., Ransom C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A."
RL J. Bacteriol. 184:3086-3095(2002).
DR EMBL; AL672113; CAD31586.1; -.
DR HSP; P77407; IPOY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
SQ SEQUENCE 389 AA; 42703 MW; 6678D2C96A7E5204 CRC64;

Query Match
Best Local Similarity 59.4%; Score 50.5; DB 2; Length 389;
Matches 9; Conservative 42.9%; Pred. No. 3.4;
Mismatch 2; Indels 7; Gaps 1;

QY 1 CADGPTL-----REWISFC 14
DB 243 CADGKEVIFSVQNDREWNFC 257

RESULT 4
Q7OC63 PRELIMINARY; PRT; 1123 AA.
ID Q7OC63;
AC Q7OC63;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE AgCP1221.
GN Name=agCG53078; ORFNames=ENSANG0000018866;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA801008859; EAA08177.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.

```

DR GO; GO:0008898; F:homocysteine S-methyltransferase activity; IEA.  
 DR GO; GO:0004672; F:protein kinase activity; IEA.  
 DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot. Kinase.  
 DR InterPro; IPR003726; S\_methyl\_trans.  
 DR Pfam; PF00069; PKinase; 1.  
 DR Pfam; PF02574; S\_methyl\_trans; 1.  
 DR ProDom; PD000001; Prot. Kinase; 1.  
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 SO SEQUENCE 1123 AA; 12006 MW; D3CC001D8D482AF CRC64;

Query Match 56.5%; Score 48; DB 2; Length 1123;  
 Best Local Similarity 75.0%; Pred. No. 29;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ADGPTLRWISF 13  
 Db 969 ADHPVTRWISF 980

RESULT 5  
 ID 082CW2 PRELIMINARY; PRT; 283 AA.

AC 082CW2; 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE Putative ICLR-family transcriptional regulator.  
 OS OrderedLocustNames=SAV5226;  
 GN Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=33903;  
 RX SEQUENCE FROM N.A.  
 RP STRAIN=MA-4680;  
 RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinozaki M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinozaki M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.,  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005042; BAC72938.1; -  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006355; P:regulation of transcription; IEA.  
 DR InterPro; IPR005471; HTH\_ICLR.  
 DR InterPro; IPR009058; wing\_hlx\_DNA\_bnd.  
 DR Pfam; PF01614; ICLR; 1.  
 DR Complete proteome.  
 SO SEQUENCE 283 AA; 30503 MW; F63B1705578EE67 CRC64;

Query Match 55.9%; Score 47.5; DB 2; Length 283;  
 Best Local Similarity 50.0%; Pred. No. 8.1;  
 Matches 8; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 CADGPTLRWISF 13  
 Db 152 CABGPTLRWISF 167

RESULT 6

ORRKMS  
 ID 09RKMS PRELIMINARY; PRT; 319 AA.  
 AC 09RKMS; 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE Putative Merf family transcriptional regulator.  
 GN ORFNames=SCD17.06c;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1902;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=AS(2) / M145;  
 RX MEDLINE=21996410; PubMed=1200953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Klee H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinowitz B., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- SIMILARITY: Contains 1 HTH merf-type DNA-binding domain.  
 DR EMBL; AL939118; CAB56383.1; -  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription; IEA.  
 DR InterPro; IPR00551; HTH\_Merf.  
 DR InterPro; IPR009061; Putativ\_DNA\_bind.  
 DR Pfam; PF00376; Merf; 1.  
 DR PRINTS; PR00040; HTHMERF.  
 DR SMART; SM00422; HTH\_MERF.1.  
 DR PROSITE; PSS0937; HTH\_MERF\_2; 1.  
 KW Complete proteome; DNA-binding.  
 SO SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;

Query Match 54.1%; Score 46; DB 2; Length 319;  
 Best Local Similarity 70.0%; Pred. No. 17;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DGPTLRWIS 12  
 Db 258 DGPTLRWIS 267

RESULT 7  
 ID 07PP6 PRELIMINARY; PRT; 347 AA.  
 AC 07PP6; 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE ENSANGP0000020769 (Fragment).  
 GN Name=ENSANG00000018280;  
 OS Anopheles gambiae str. FEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008944; EAA10075.2; -



```
DR GO; GO:0008896; F:homocysteine S-methyltransferase activity; IEA.
DR InterPro; IPR003726; S_methyl_trans.
DR Pfam; PF02574; S-methyl_trans; 1.
FT NON TER
SQ SEQUENCE 347 AA; 38585 MW; 66FF58A1000CDA4F CRC64;

Query Match 54.1%; Score 46; DB 2; Length 347;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CADGPTLRWISF 13
Db 201 CDEYTVRWFISF 213

RESULT 8
O7RUAS PRELIMINARY; PRT; 108 AA.
AC O7RUAS;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DE Hypothetical protein B24B19.30.
GN Name=NCU03933.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetiales; Sordariales; Sordariaceae; Neurospora.
OC NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR7A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehan B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankov P., Pedersen D., Nelson M., Washburne M.,
RA Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseis M., Mauceli E., Bielke C., Rudd S., Frisman D.,
RA Kryzocova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Marino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA Desguza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Nativ D.O., Alex L.A., Mannheim G., Ebdole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABX01000719; EAA28336.1; -.
DR Hypothetical protein.
KW SEQUENCE 108 AA; 11994 MW; 093DC0D9617A252E CRC64;

Query Match 52.9%; Score 45; DB 2; Length 108;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CADGPTLRWISF 14
Db 70 CQCPILRWISWMC 83

RESULT 9
O8C4M6 PRELIMINARY; PRT; 173 AA.
AC O8C4M6;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
DE library, clone:Cl30070D15 product:unclassifiable, full insert
DE sequence.
```

```
GN Name=Cl30070B15Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiji K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Atzawa K., Akiyama T., Arahata T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hasegaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata H., Konda M., Koya S.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Nakamura Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaru-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK081706; BAC38302.1; -.
DR MGI; MGI:2444974; Cl30070B15Rik.
SQ SEQUENCE 173 AA; 19340 MW; 6227DD6725E52FCD CRC64;

Query Match 53.9%; Score 45; DB 2; Length 173;
Best Local Similarity 53.6%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```



QY 4 GPTIREWISFC 14  
|:|:|:|:  
DB 75 GVTIREWASWC 85

## RESULT 10

OG6N1X5 PRELIMINARY; PRT; 209 AA.  
AC OG6N1X5;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=RP44277;  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiales; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CGA009 / ATCC BAA-98;  
RX PubMed=14704707; DOI=10.1038/nbt1923;  
RA Lattimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,  
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,  
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y Torres J.L., Perez C.,  
RA Harrison C.H., Gibson J., Harwood C.S.;  
RT "Complete genome sequence of the metabolically versatile  
RT photosynthetic bacterium Rhodopseudomonas palustris.";  
RL Nat. Biotechnol. 22:55-61(2004).

DR EMBL; BX572606; CAB29718.1; -  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR003571; HEAT.  
DR Pfam; PF02985; HEAT; 2.

KM Complete proteome; Hypothetical protein.  
SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84DB040EE CRC64;

Query Match 52.9%; Score 45; DB 2; Length 209;  
Best Local Similarity 50.0%; Pred. No. 16;  
Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 1 CADG-----PTIREWIS 12  
|:|:|:|:|:  
DB 98 CADTGCEAALPTIREWIS 115

## RESULT 11

OG8XZNS PRELIMINARY; PRT; 309 AA.

AC OG8XZNS;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 26, Last annotation update)  
DE PROBABLE TRANSCRIPTION REGULATOR PROTEIN.

GN Name=RS04642; OrderedLocustNames=RScl360;  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;  
RA Saranobat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Broctier P., Camus J.C., Cattolico L.,  
RA Chandler M., Choise N., Chandel-Renard C., Cunac S., Demange N.,  
RA Gaopin C., Lavie M., Moisan A., Robert C., Saurin W., Schiek T.,  
RA Signier F., Thebault P., Whalen W., Wincker P., Levy M.,  
RA Weissbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
RL Nature 415:497-502(2002).

DR EMBL; AL646064; CAD15062.1; -  
DR HSP; O9WKC7; 11XC.  
DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR Pfam; PF00126; HTH 1; 1.  
DR Pfam; PF03466; LysR\_substrate; 1.  
DR PROSITE; PS50931; HTH\_LYSR; 1.  
KM Complete proteome.

SQ SEQUENCE 309 AA; 33774 MW; 73351741CE83182 CRC64;

Query Match 52.9%; Score 45; DB 2; Length 309;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CADGPTIREW 10  
|:|:|:|:  
DB 221 CTDCAVIREW 230

## RESULT 12

OG9P858 PRELIMINARY; PRT; 443 AA.

AC OG9P858;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein.

OS Phaeosphaeria nodorum (Septoria nodorum).  
OC Plasmid pBla1.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;  
OC Pleosporales; Phaeosphaeriaceae; Phaeosphaeria.  
OX NCBI\_TaxID=13684;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=BS444;  
RA Rawson J.M.;  
RT "Transposable elements in the phytopathogenic fungus Stegomyces  
RT nodorum.";  
RL Thesis (2000), PhD thesis, University of Birmingham, UK.

RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=BS444;  
RA Rawson J.M., Cutler S.B., Caten C.E.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ277966; CAB91876.1; -  
KM Hypothetical protein; Plasmid.  
SQ SEQUENCE 443 AA; 49466 MW; 367E0762EB39E68 CRC64;

Query Match 52.9%; Score 45; DB 2; Length 443;  
Best Local Similarity 58.3%; Pred. No. 36;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CADGPTIREWIS 12  
|:|:|:|:|:  
DB 170 CSENGTIREWIT 181

## RESULT 13

OG6QHD2 PRELIMINARY; PRT; 173 AA.

AC OG6QHD2;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE P32 (Fragment).

OS Gallid herpesvirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Iltovirus.  
OX NCBI\_TaxID=10386;

RN [1]  
RP SEQUENCE FROM N.A.

RC Villareal L.Y., Brandão P.E., Ferreira A.P., Doretto L.J.,  
RA D'elboux A.N.;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY541676; AAG48543.1; -  
DR InterPro; IPR003363; Herpes\_GC.

DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF02400; Herpes\_gg; 1.  
 FT NON\_TER 1  
 FT NON\_TER 173 173  
 SQ SEQUENCE 173 AA; 19130 MW; 5A64A1956CEB9B13 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 173;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 DB 120 CLDMPPLRPWTTC 133

## RESULT 14

ID 06PL14 PRELIMINARY; PRT; 178 AA.

AC 06PL14;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE P32 (Fragment)  
 OS Gallid herpesvirus 1.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Iltovirus.  
 NCBI\_TaxId=10386;

RP SEQUENCE FROM N.A.  
 RA Villarreal L.Y., Ferreira P.E., Peguini M.R., Ito N.M., Gama N.,  
 RA Ferreira C.A., Ferreira A.J.,  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY598339; AA09767.1; -  
 DR InterPro; IPR003363; Herpes\_gg.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF02400; Herpes\_gg; 1.  
 FT NON\_TER 1  
 FT NON\_TER 178 178  
 SQ SEQUENCE 178 AA; 19896 MW; 6CC47102594537EF CRC64;

Query Match 51.8%; Score 44; DB 2; Length 178;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 DB 120 CLDMPPLRPWTTC 133

## RESULT 15

ID 09L059 PRELIMINARY; PRT; 209 AA.

AC 09L059;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE Hypochemical protein SC02976.  
 GN ORFNames=SC50.04c;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NCBI\_TaxId=1902;

RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieiser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Frazer A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabbittowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.,  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).",  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL39114; CAB87326.1; -  
 KW Complete proteome; Hypochemical protein.  
 SQ SEQUENCE 209 AA; 24308 MW; 34BAF6CA5D96AB7 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 209;  
 Best Local Similarity 70.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CADGPTLRW 10  
 DB 66 CAQGPALRYW 75

## RESULT 16

ID 067642 PRELIMINARY; PRT; 292 AA.

AC 067642;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE Glycoprotein G.  
 OS Gallid herpesvirus 1.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Iltovirus.  
 NCBI\_TaxId=10386;

RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA challenge strain;  
 RA MEDLINE=97033380; PubMed=8879127;  
 RX Wild M.A., Cook S., Cochran M.,  
 RT "A genomic map of infectious laryngotracheitis virus and the sequence  
 RT and organization of genes present in the unique short and flanking  
 RT regions.",  
 RT Virus Genes 12:107-116(1996).  
 DR EMBL; U28832; AAC55098.1; -  
 DR InterPro; IPR003363; Herpes\_gg.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF02400; Herpes\_gg; 1.  
 SQ SEQUENCE 292 AA; 31696 MW; 7B2D3D35D9F32E8 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 292;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 DB 229 CLDMPPLRPWTTC 242

## RESULT 17

ID 086553 PRELIMINARY; PRT; 298 AA.

AC 086553;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE P32 protein.  
 GN Name=P32;  
 OS Gallid herpesvirus 1.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Iltovirus.  
 NCBI\_TaxId=10386;

RP SEQUENCE FROM N.A.  
 RP MEDLINE=94025939; PubMed=8212855; DOI=10.1016/0168-1702(93)90054-Q;  
 RX Kongwan K., Johnson M.A., Pridoux C.T., Sheppard M.,  
 RT "Identification of an infectious laryngotracheitis virus gene encoding  
 RT an immunogenic protein with a predicted M(r) of 32 kilodaltons.";

RL Virus Rec. 29:125-140(1993).  
 DR EMBL; S66009; AAB28457.1; -.  
 DR InterPro; IPR003363; Herpes\_gg.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF02400; Herpes\_gg; 1.  
 SQ SEQUENCE 298 AA; 32325 MW; 737E428E3CBA4215 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 298;  
 Best Local Similarity 50.0%; Pred. No. 35;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CADGPTLRWISFC 14  
 DB 228 CLDMPPLRPWTVC 241

RESULT 18  
 PHS2\_SOLUT STANDARD; PRT; 974 AA.  
 ID PHS2\_SOLUT  
 AC P53535;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alpha-1,4 glucan phosphorylase, L-2 isozyme, chloroplast precursor  
 DE (EC 2.4.1.1) (Starch phosphorylase L-2).  
 GN Name-STP-1;  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 NCBI\_TaxID=4113;  
 RX MEDLINE=95201249; PubMed=7894019;  
 RA Somerville U., Basner A., Greve B., Steup M.;  
 RT "A second L-type isozyme of potato glucan phosphorylase: cloning,  
 RT antisense inhibition and expression analysis."  
 RL Plant Mol. Biol. 27:567-576(1995).  
 CC -1- FUNCTION: Phosphorylase is an important allosteric enzyme in  
 CC carbohydrate metabolism. Enzymes from different sources differ in  
 CC their regulatory mechanisms and in their natural substrates.  
 CC However, all known phosphorylases share catalytic and structural  
 CC properties.  
 CC -1- CATALYTIC ACTIVITY: {(1,4)-alpha-D-glucosyl}(N) + phosphate =  
 CC {(1,4)-alpha-D-glucosyl}(N-1) + alpha-D-glucose 1-phosphate.  
 CC -1- COFACTOR: Pyridoxal phosphate.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast; amyloplast.  
 CC -1- TISSUE SPECIFICITY: Leaves.  
 CC -1- SIMILARITY: Belongs to the glycogen phosphorylase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X73684; CAA52036.1; -.  
 DR PIR; S53489; S34189.  
 DR HSPD; P06738; 1YGP.  
 DR InterPro; IPR00811; Glyco\_trans\_35.  
 DR Pfam; PF00343; Phosphorylase; 1.  
 DR PROSITE; PS00102; PHOSPHORYLASE; 1.  
 DR PROSITE; PS00102; PHOSPHORYLASE; 1.  
 KW Allosteric enzyme; Amyloplast; Carbohydrate metabolism; Chloroplast;  
 KW Glycosyltransferase; Multigene family; Pyridoxal phosphate;  
 KW Transferase; Transit peptide.  
 FT TRANSIT 1 81  
 FT CHAIN 82 974 Alpha-1,4 glucan phosphorylase, L-2  
 FT BINDING 820 820 Isozyme.  
 FT BINDING 820 820 Pyridoxal phosphate (By similarity).

SQ SEQUENCE 974 AA; 110700 MW; 5EF8A23C237463D8 CRC64;

Query Match 51.8%; Score 44; DB 1; Length 974;  
 Best Local Similarity 58.3%; Pred. No. 1,2e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DGPTLRWISFC 14  
 DB 619 NGVTPRRWISFC 630

RESULT 19  
 ID OGB126 PRELIMINARY; PRT; 997 AA.  
 AC OGB126;  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Similar to CA61451|IP1869 Candida albicans IP1869.  
 GN ORFNames=DEHA00G147959;  
 OS Debaryomyces hansenii CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
 NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marcq C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boissarie A., Boyer J., Catolico L., Confariollet F., de Darvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppe A.,  
 RA Hantreya F., Hennequin C., Jaumaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicand J.M., Nikoleki M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swenne D., Tekala F., Wesolowski-Jouvet M., Westhof E., Wirth B.,  
 RA Zentou-Meyer M., Zivanovic J., Bolotin-Fukuhara M., Thierry A.,  
 RA Boucher C., Caudron B., Scarpelli C., Galliardin C., Weissenbach J.,  
 RA Winkler P., Souciet J.L.;  
 RT "Genome evolution in Yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RA Genoscope;  
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382139; CAG90631.1; -.  
 DR InterPro; IPR011046; WD40\_like.  
 SQ SEQUENCE 997 AA; 112803 MW; 3C05D6EAF05875C CRC64;  
 Query Match 51.8%; Score 44; DB 2; Length 997;  
 Best Local Similarity 87.5%; Pred. No. 1,3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 LREWISFC 14  
 DB 956 LREWISFC 963

RESULT 20  
 ID O8AY57 PRELIMINARY; PRT; 1008 AA.  
 AC O8AY57;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Sodium/potassium ATPase alpha subunit isoform 2.  
 GN Name=ATP1A2;  
 OS Fundulus heteroclitus (Killifish) (Mummichog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OC NCBI_TaxID=8078;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA PubMed=14961245;
RT Sempke J.W., Green H.J., Schulte P.M.;
RT "Molecular Cloning and Characterization of Two Na/K-ATPase Isoforms in
  Fundulus heteroclitus."
RL Mar. Biotechnol. 4:512-519(2002).
DR EMBL; AY057073; AA18003.1; -.
DR HSSP; P06685; IM07.
DR GO; GO:0005890; C:sodium:potassium-exchanging ATPase complex; ISS.
DR GO; GO:0005391; F:sodium:potassium-exchanging ATPase activity; ISS.
DR GO; GO:0015991; P:ATP hydrolysis coupled proton transport; ISS.
DR GO; GO:0030641; P:hydrogen ion homeostasis; ISS.
DR GO; GO:0006813; P:potassium ion transport; ISS.
DR GO; GO:0006814; P:sodium ion transport; ISS.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR005775; Na/K_ATPase_alph.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR TIGRAME; TIGR01106; ATPase-ITC-X-K; 1.
DR TIGRAME; TIGR01494; ATPase_P-type; 4.
DR PROSITE; PS00154; ATPASE_E1-E2; UNKNOWN 1.
SQ SEQUENCE 1008 AA; 111293 MW; EA3A/CED8E33E037 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 1008;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14
   |||||
Db 70 PTPPEWIKFC 79

RESULT 21
Q6VYM7 PRELIMINARY; PRT; 1011 AA.
ID O6VYM7;
AC O6VYM7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Na/K ATPase alpha subunit isoform 3. (Salmo gairdneri).
OS Oncorhynchus mykiss (Rainbow trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_TaxID=8022;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA PubMed=1461032;
RT "Na(+)/K(+)-ATPase alpha-isoform switching in gills of rainbow trout
  (Oncorhynchus mykiss) during salinity transfer."
RL J. Exp. Biol. 206:4475-4486(2003).
DR EMBL; AY139388; AA082787.1; -.
DR HSSP; P04191; 1KTU.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.

```

```

DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m...; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid...; IEA.
DR GO; GO:0015077; F:monovalent inorganic cation transporter act...; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0015672; P:monovalent inorganic cation transport; IEA.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR005775; Na/K_ATPase_alph.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR TIGRAME; TIGR01106; ATPase-ITC-X-K; 1.
DR PROSITE; PS00154; ATPASE_E1-E2; UNKNOWN 1.
SQ SEQUENCE 1011 AA; 111140 MW; 06D12FA68A23456C CRC64;

Query Match 51.8%; Score 44; DB 2; Length 1011;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14
   |||||
Db 72 PTPPEWIKFC 81

RESULT 22
ID ATIA TORCA STANDARD; PRT; 1022 AA.
AC P05025;
DT 13-AUG-1987 (rel. 05, Last sequence update)
DT 05-JUL-2004 (rel. 44, Last annotation update)
DE Sodium/potassium-transporting ATPase alpha chain precursor
DE (EC 3.6.3.9) (Sodium pump alpha chain) (Na+/K+ ATPase alpha chain).
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hyngnosquales; Pristiogaster; Batoidae;
OC Torpediniformes; Torpedinidae; Torpedo.
OC NCBI_TaxID=7787;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=85296307; PubMed=2993905;
RA Kawakami K., Noguchi S., Noda M., Takahashi H., Ohra T., Kawamura M.,
RA Nojima H., Nagano K., Hirose T., Inayama S., Hayashida H., Miyata T.,
RA Numa S.;
RT "Primary structure of the alpha-subunit of Torpedo californica (Na+ +
  K+)ATPase deduced from cDNA sequence."
RL Nature 316:733-736(1985).
CC -1- FUNCTION: This is the catalytic component of the active enzyme,
  which catalyzes the hydrolysis of ATP coupled with the exchange of
  sodium and potassium ions across the plasma membrane. This action
  creates the electrochemical gradient of sodium and potassium ions,
  providing the energy for active transport of various nutrients.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+)(In) + K(+)(Out) = ADP +
  phosphate + Na(+)(Out) + K(+)(In).
CC -1- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and
  gamma.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
  ATPases). Subfamily IIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X02810; CAA26578.1; -.

DR PIR; S00503; S00503.

DR HSSP; P06685; 1M07.

DR InterPro; IPR001757; ATPase\_E1-E2.

DR InterPro; IPR006069; Cation\_ATPase.

DR InterPro; IPR006068; Cation\_ATPase\_C.

DR InterPro; IPR004014; Cation\_ATPase\_N.

DR InterPro; IPR005834; Dehalo-like\_hydro.

DR InterPro; IPR008250; E1-E2\_ATPase\_reg.

DR InterPro; IPR005775; Na/K\_ATPase\_alph.

DR Pfam; PF00689; Cation\_ATPase\_C; 1.

DR Pfam; PF00690; Cation\_ATPase\_N; 1.

DR Pfam; PF00122; E1-E2\_ATPase; 1.

DR Pfam; PF00702; Hydrolyase; 1.

DR PRINTS; PR00119; CATATPASE.

DR PRINTS; PR00121; NAKATPASE.

DR TIGRFAm; TIGR01106; ATPase-11C-X-K; 1.

DR TIGRFAm; TIGR01494; ATPase\_P-type; 4.

DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.

KM APP-binding; Direct protein sequencing; Hydrolyase; Phosphorylation;

KM Sodium/potassium transport; Transmembrane.

PT PROSP 1 5

FT CHAIN 6 1022

FT DOMAIN 6 87 Sodium/potassium-transporting ATPase

FT TRANSSEM 88 108 Alpha chain.

FT DOMAIN 109 131 Cytoplasmic (Potential).

FT TRANSSEM 132 152 Potential.

FT DOMAIN 153 288 Potential.

FT TRANSSEM 289 320 Cytoplasmic (Potential).

FT TRANSSEM 321 338 Potential.

FT TRANSSEM 339 771 Potential.

FT TRANSSEM 772 791 Potential.

FT TRANSSEM 792 801 Potential.

FT TRANSSEM 802 822 Potential.

FT TRANSSEM 823 842 Cytoplasmic (Potential).

FT TRANSSEM 843 865 Potential.

FT TRANSSEM 866 917 Potential.

FT TRANSSEM 918 937 Potential.

FT TRANSSEM 938 950 Cytoplasmic (Potential).

FT TRANSSEM 951 969 Potential.

FT TRANSSEM 970 984 Potential.

FT TRANSSEM 985 1005 Potential.

FT TRANSSEM 1006 1022 Potential.

FT MOD RES 16 16 Cytoplasmic (Potential).

FT ACT\_SITE 376 376 Phosphoserine (by PKC) (By similarity).

FT MOD RES 942 942 4-aspartylphosphate intermediate (By similarity).

FT BINDING 82 84 Phosphoserine (by PKA) (By similarity).

FT METAL 716 716 Binding of phosphoinositide-3 kinase (By similarity).

FT METAL 720 720 Magnesium (By similarity).

FT METAL 720 720 Magnesium (By similarity).

SO SEQUENCE 1022 AA; 112429 MW; D92FE737847D73C2 CRC64;

Query Match 51.8%; Score 44; DB 1; Length 1022;

Best Local Similarity 70.0%; Pred. No. 1.3e+00; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 3;

Qy 5 PTLREWISFC 14

Db 84 PTPPEWIKFC 93

DT 25-JAN-2005 (Rel. 46, Last annotation update)

DE Sodium/potassium-transporting ATPase alpha-1 chain precursor

DR (EC 3.6.3.9) (sodium pump 1) (Na<sup>+</sup>/K<sup>+</sup> ATPase 1).

GN Name=ATP1A1;

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RX MEDLINE=67057096; PubMed=2430951;

RA Kawakami K., Ohta T., Nijima H., Nagano K.;

RT "Primary structure of the alpha-subunit of human Na,K-ATPase deduced from cDNA sequence.";

RT J. Biochem. 100:389-397 (1986).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RC TISSUE=Retinal pigment epithelium;

RX MEDLINE=95237606; PubMed=7536695; DOI=10.1016/0378-1119(94)00812-7;

RA Ruiz A., Bhat S.P., Bok D.;

RT "Characterization and quantification of full-length and truncated Na,K-ATPase alpha 1 and beta 1 RNA transcripts expressed in human retinal pigment epithelium.";

RT Gene 155:179-184 (1995).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RC TISSUE=Brain, Cervix, and Skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin C.M., Hong L.,

RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udell T.B., Toshylyuk S., Carninci P., Prange C.,

RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,

RA Scherch A., Schein J.E., Jones S.U.M., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [4]

RP SEQUENCE OF 471-619 FROM N.A.

RA Ovchinnikov Y.A., Monastyrskaya G.S., Arsenyan S.G., Brode N.E.,

RA Petrunkin K.E., Grishin A.V., Arzamazova N.M., Severtsova I.V.,

RA Modyanov N.N.;

RT "Amino acid sequence of the 17-kilodalton fragment of the cytoplasmic region of the alpha-subunit of Na<sup>+</sup>/K<sup>+</sup>-ATPase.";

RT Dokl. Biochem. 288:270-272 (1986).

RN [5]

RP SEQUENCE OF 253-341 AND 420-444 FROM N.A.

RX MEDLINE=87247232; PubMed=3036582; DOI=10.1016/0014-5793(87)80677-4;

RA Sverdlov G.D., Monastyrskaya G.S., Brode N.E., Unsharyov Y.A.,

RA Alilikmet R.L., Melkov A.M., Smirnov Y.V., Malyshev I.V.,

RA Dulubova I.E., Petrunkin K.E., Grishin A.V., Kiyetkin N.I.,

RA Kostina M.B., Sverdlov V.E., Modyanov N.N., Ovchinnikov Y.A.;

RT "The family of human Na<sup>+</sup>/K<sup>+</sup>-ATPase genes. No less than five genes and/or pseudogenes related to the alpha-subunit.";

RT FEBS Lett. 217:275-276 (1987).

RN [6]

RP SEQUENCE OF 198-943 FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=88068506; PubMed=2891135;

RA Chehab F.F., Kan Y.W., Law M.L., Hartz J., Kao F.T., Blostein R.;

RT "Human placental Na<sup>+</sup>/K<sup>+</sup>-ATPase alpha subunit: cDNA cloning, tissue expression, DNA polymorphism, and chromosomal localization.";



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueffing T.B., Tothiyuki S., Canninci P., Prange C.,  
 RA Raha S.S., Locnellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Kravinsky M.I., Skalska U., Smallus D.E., Schermer A., Schein J.E.,  
 RA Jones S.J., Matra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Embryo;  
 RC MEDLINE=22311132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.,  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 initiative";  
 RT Dev. Dyn. 225:384-391(2002).  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Embryo;  
 RC Klein S., Strausberg R.,  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC043743; AAA43743.1; -  
 DR HSSP; P06685; IM07.  
 DR GO; GO:0016021; C:Integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0016787; F:Hydrolase activity; IEA.  
 DR GO; GO:0016820; F:Hydrolase activity, acting on acid anhydrid. . .; IEA.  
 DR GO; GO:0015077; F:monovalent inorganic cation transporter act. . .; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR GO; GO:0015672; P:monovalent inorganic cation transport; IEA.  
 DR InterPro; IPR001757; ATPase\_E1-E2.  
 DR InterPro; IPR006069; Cation ATPase.  
 DR InterPro; IPR006068; Cation ATPase\_C.  
 DR InterPro; IPR004014; Cation ATPase\_N.  
 DR InterPro; IPR005834; Dehal like hydro.  
 DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
 DR InterPro; IPR005775; Na/K\_ATPase\_alph.  
 DR Pfam; PF00669; Cation\_ATPase\_C; 1.  
 DR Pfam; PF00122; E1-E2\_ATPase\_N; 1.  
 DR Pfam; PF00702; Hydrolase; 1.  
 DR PRINTS; PR00119; CATATPASE.  
 DR PRINTS; PR00121; NAKATPASE.  
 DR TIGRfams; TIGR01106; ATPase-ITC-X-K; 1.  
 DR TIGRfams; TIGR01494; ATPase\_P-type; 4.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; UNKNOWN 1.  
 SQ SEQUENCE 1025 AA; 112954 MW; FAOC021193F5288E CRC64;

RESULT 25  
 ID A1A4 RAT STANDARD; PRT; 1028 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE (Sodium/potassium-transporting ATPase alpha-4 chain (EC 3.6.3.9)  
 DE (Sodium pump 4) (Na+/K+ ATPase 4).  
 GN Name=Atpla4; Synonyms=Atpla2;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;  
 RX MEDLINE=95108076; PubMed=7809153;  
 RA Shamraj O.I., Lingrel J.B.,  
 RT "A putative fourth Na+,K(+)-ATPase alpha-subunit gene is expressed in  
 testis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12952-12956(1994).  
 CC -1- FUNCTION: This is the catalytic component of the active enzyme,  
 which catalyzes the hydrolysis of ATP coupled with the exchange of  
 sodium and potassium ions across the plasma membrane. This action  
 creates the electrochemical gradient of sodium and potassium ions,  
 providing the energy for active transport of various nutrients.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+)(In) + K(+)(Out) = ADP +  
 phosphate + Na(+)(Out) + K(+)(In).  
 CC -1- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and  
 gamma.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type  
 ATPases). Subfamily IIC.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1eb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; U15176; AAB81285.1; -  
 DR HSSP; P06685; IM07.  
 DR RGD; 61952; Atpla4.  
 DR GO; GO:0005890; C:sodium/potassium-exchanging ATPase complex; ISS.  
 DR GO; GO:0005391; F:sodium/potassium-exchanging ATPase activity; ISS.  
 DR GO; GO:0015991; P:ATP hydrolysis coupled proton transport; ISS.  
 DR GO; GO:0030641; P:hydrogen ion homeostasis; ISS.  
 DR GO; GO:0006813; P:potassium ion transport; ISS.  
 DR GO; GO:0006814; P:sodium ion transport; ISS.  
 DR GO; GO:0030317; P:sperm motility; ISS.  
 DR InterPro; IPR001757; ATPase\_E1-E2.  
 DR InterPro; IPR006069; Cation ATPase.  
 DR InterPro; IPR006068; Cation ATPase\_C.  
 DR InterPro; IPR004014; Cation ATPase\_N.  
 DR InterPro; IPR005834; Dehal like hydro.  
 DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
 DR InterPro; IPR005775; Na/K\_ATPase\_alph.  
 DR Pfam; PF00669; Cation\_ATPase\_C; 1.  
 DR Pfam; PF00122; E1-E2\_ATPase\_N; 1.  
 DR Pfam; PF00702; Hydrolase; 1.  
 DR PRINTS; PR00119; CATATPASE.  
 DR PRINTS; PR00121; NAKATPASE.  
 DR TIGRfams; TIGR01106; ATPase-ITC-X-K; 1.  
 DR TIGRfams; TIGR01494; ATPase\_P-type; 4.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 KW ATP-binding; Hydrolase; Magnesium; Metal-binding; Multigene family;  
 phosphorylation; Sodium/potassium transport; Transmembrane.

QY 5 PTLREWSIFC 14  
 DB 86 PTLPEWIKFC 95

Query Match 51.8%; Score 44; DB 2; Length 1025;  
 Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;



```

FT DOMAIN 1 92 Cytoplasmic (Potential).
FT TRANSMEM 93 113 Potential.
FT DOMAIN 114 137 Luminal (Potential).
FT TRANSMEM 138 158 Potential.
FT DOMAIN 159 294 Cytoplasmic (Potential).
FT TRANSMEM 295 314 Potential.
FT DOMAIN 315 326 Luminal (Potential).
FT TRANSMEM 327 344 Potential.
FT DOMAIN 345 777 Cytoplasmic (Potential).
FT TRANSMEM 778 797 Potential.
FT DOMAIN 798 807 Luminal (Potential).
FT TRANSMEM 808 828 Potential.
FT DOMAIN 829 848 Cytoplasmic (Potential).
FT TRANSMEM 849 871 Potential.
FT DOMAIN 872 923 Luminal (Potential).
FT TRANSMEM 924 943 Potential.
FT TRANSMEM 944 956 Cytoplasmic (Potential).
FT TRANSMEM 957 975 Potential.
FT DOMAIN 976 990 Luminal (Potential).
FT TRANSMEM 991 1011 Potential.
FT DOMAIN 1012 1028 Cytoplasmic (Potential).
FT ACT_SITE 382 382 4-aspartylphosphate intermediate (By similarity).
FT MOD_RES 948 948 Phosphoserine (by PKA) (By similarity).
FT BINDING 87 89 Binding of phosphoinositide-3 kinase (By similarity).
FT METAL 722 722 Magnesium (By similarity).
FT METAL 726 726 Magnesium (By similarity).
SQ SEQUENCE 1028 AA; 114004 MW; 958FE008735D06FA CRC64;

```

```

Query Match
Best Local Similarity 51.8%; Score 44; DB 1; Length 1028;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 5 PTLREWISFC 14
Db 89 PTPPWIKFC 98

```

```

RESULT 26
Q9KIE9 PRELIMINARY; PRT; 405 AA.
AC Q9KIE9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE PKB.
GN Name=fkBE;
OS Streptomyces hygroscopicus subsp. ascomyceticus;
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=132248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20323220; PubMed=10863099; DOI=10.1016/S0378-1119(00)00171-2;
RA Wu K., Chung L., Revelli W.P., Katz L., Reeves C.D.;
RT "The FK520 gene cluster of Streptomyces hygroscopicus var.
RT ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual
RT polycyclic extender units."
RL Gene 251:81-90(2000).
DR EMBL; AF235504; AAF86384.1; -.
DR HSSP; P77407; 1PQY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
SQ SEQUENCE 405 AA; 43696 MW; DC2569DFC914AD6F CRC64;

```

```

Query Match
Best Local Similarity 51.2%; Score 43.5; DB 2; Length 405;
Matches 9; Conservative 1; Mismatches 2; Indels 7; Gaps 1;
Qy 3 DGPTL-----REWISFC 14
|||: |||||

```

```

Db 252 DGQTNLGLQNEREMWASFC 270

```

```

RESULT 27
Q9NEX6 PRELIMINARY; PRT; 934 AA.
AC Q9NEX6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein Y105E8A.21.
GN ORFNames=Y105E8A.21;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Felodertidae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sulston J.E.;
RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL132876; CAC48140.1; -.
DR WormBase; WBGene00013679; Y105E8A.21.
DR WormPeP; Y105E8A.21; CE25162.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
KW Hypothetical protein.
SQ SEQUENCE 934 AA; 104855 MW; SED4ELD03DB06F24 CRC64;

```

```

Query Match
Best Local Similarity 51.2%; Score 43.5; DB 2; Length 934;
Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 1 CADGPTLREW-ISF 13
Db 899 CVDGTTSRDWPVSF 912

```

```

RESULT 28
Q9N0Z5 PRELIMINARY; PRT; 127 AA.
AC Q9N0Z5;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Na/K ATPase alpha 2 subunit (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21600302; PubMed=11738066; DOI=10.1016/S0008-6363(01)00412-6;
RA Fransen P., Hendrickx U., Brutsaert D.L., Sys S.U.;
RT "Distribution and role of Na(+)/K(+) ATPase in endocardial
RT endothelium."
RL Cardiovasc. Res. 52:487-499(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Hendrickx U., Fransen P.;
RL Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF235025; AAF60311.2; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.

```



```

DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR004014; Cation_ATPase_N.
DR Pfam; PF00690; Cation_ATPase_N; 1.
FT NON_TER 1 127 127
SQ SEQUENCE 127 AA; 14082 MW; 301B90F956954550 CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 127;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14
DB 104 PTPPEWVKFC 113

RESULT 29
O8HYW6 PRELIMINARY; PRT; 171 AA.
AC O8HYW6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sodium/potassium-transporting ATPase alpha-3 chain (EC 3.6.3.9)
DE (Fragment).
GN Name=atp1a3;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal medulla;
RA Benavides A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ496458; CAD4286.1; -.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0016820; F: hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0005911; F: sodium/potassium-exchanging ATPase activity; IEA.
DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR PRINTS; PR00121; NAKATPASS.
DR Hydrolase.
FT NON_TER 1 171 171
SQ SEQUENCE 171 AA; 19015 MW; B61570772C03945A CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 171;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14
DB 73 PTPPEWVKFC 82

RESULT 30
O866A9 PRELIMINARY; PRT; 176 AA.
AC O866A9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Alpha subunit of equine Na/K ATPase (fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Budik S.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ539381; CAD62375.1; -.
DR HSSP; P08515; 1BGS.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0015662; F: ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0016820; F: hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR PRINTS; PR00121; NAKATPASS.
FT NON_TER 1 176 176
SQ SEQUENCE 176 AA; 19615 MW; C5F349EBB74444A8 CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 176;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLREWISFC 14
DB 68 PTPPEWVKFC 77

RESULT 31
O9M060 PRELIMINARY; PRT; 245 AA.
AC O9M060;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 6 (EIF-6)-like protein
DE (A1355620).
GN Name=Flitf 30; Synonyms=At3g55620;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmach E., Dzonek H., Anegoe W., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Shim P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai U., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinzaki K.,
RA Becker J., Theologis A., Davis R.W.;

```

```

RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL161667; CAB81587.1; -.
DR EMBL; BT009656; AAM75806.1; -.
DR EMBL; AY128351; AAM91554.1; -.
DR PIR; T47701; T47701.
DR HSSP; O12522; 1G62.
DR GO; GO:0003744; P:translation initiation factor activity; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR InterPro; IPR002769; eIF6.
DR Pfam; PF01912; eIF-6; 1.
DR ProDom; PD006880; eIF6; 1.
DR SMART; SM00654; eIF6; 1.
DR TIGRfam; TIGR00323; eIF-6; 1.
DR Initiation Factor.
SQ SEQUENCE 245 AA; 26482 MW; 73369A2A657F390D CRC64;

Query Match 50.6%; Score 43; DB 2; Length 245;
Best Local Similarity 53.8%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ADGPTLRWISFC 14
DB 194 AAGMTVNDWTSFC 206

RESULT 32
Q6NMU4 PRELIMINARY; PRT; 407 AA.
AC Q6NMU4;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE AT16019D.
GN Name=CG15483;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Ceiniker S.;
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; BT011561; AAS15697.1; -.
DR InterPro; IPR000345; CyC heme BS.
DR PROSITE; PS00190; CYTOCHROME C1 UNKNOWN 1.
SQ SEQUENCE 407 AA; 46390 MW; BFA2D74079EE09 CRC64;

Query Match 50.6%; Score 43; DB 2; Length 407;
Best Local Similarity 58.3%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CADGPTLRWIS 12
DB 280 CHRGPNLEWIN 291

RESULT 33
Q9VK55 PRELIMINARY; PRT; 407 AA.
AC Q9VK55;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TREMBLrel. 26, Last annotation update)
DE CG15483-PA.
GN ORFNames=CG15483;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abiri J.F., Agbayant A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockler P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusken D.R., Pacleb J.M.,
RA Palazolo M., Plittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier F., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgerl, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hogson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."

```

```

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003637; AAF53224.1; -.
DR FlyBase; FBgn0032457; CG15483.
DR InterPro; IPR000345; Cytochrome BS.
DR PROSITE; PS00190; CYTOCHROME_C_1 UNCOMMON.
SQ SEQUENCE 407 AA; 46420 MW; BFAC2105079EE508 CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 407;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CADPTLRWIS 12
DB 280 CHRGPILERMIN 291

RESULT 34
Q37839 PRELIMINARY; PRT; 469 AA.
AC Q37839;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF469 protein.
GN Name=ORF469;
OS Bacteriophage R4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=10732;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsumura M., Noguchi T., Aida T., Asayama M., Takahashi H., Shirai M.;
RT "A gene essential for the site-specific excision of actinophage R4 prophage genome from the chromosome of a lysogen.";
RL J. Gen. Appl. Microbiol. 41:53-61(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96236063; PubMed=8655526;
RA Matsumura M., Noguchi T., Yamaguchi D., Aida T., Asayama M., Takahashi H., Shirai M.;
RT "The arg gene (ORF469) encodes a site-specific recombinase responsible for integration of the R4 phage genome.";
RL J. Bacteriol. 178:3374-3376(1996).
DR EMBL; D38173; BAA07372.1; -.
DR GO; GO:0000150; P:recombinase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR011109; Recombinase.
DR InterPro; IPR006119; Recombinase.
DR Pfam; PF07508; Recombinase; 1.
DR Pfam; PF00239; Resolvase; 1.
SQ SEQUENCE 469 AA; 50656 MW; B3C37E3E2A43853C CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 469;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLRWISFC 14
DB 429 PTRRAWVDFC 438

RESULT 35
Q04270 PRELIMINARY; PRT; 490 AA.
AC Q04270;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)

```

```

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphatidylinositol 3-kinase (Fragment).
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CW-15;
RX MEDLINE=98281574; PubMed=9620264; DOI=10.1023/A:1005973423723;
RA Molendijk A.J., Irvine R.F.;
RT "Inositide signalling in Chlamydomonas: characterization of a phosphatidylinositol 3-kinase gene.";
RL Plant Mol. Biol. 37:53-66(1998).
DR EMBL; U97663; AAC50018.1; -.
DR PIR; T09084; T09084.
DR GO; GO:0005942; C:phosphoinositide 3-kinase complex; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016303; F:phosphatidylinositol 3-kinase activity; IEA.
DR InterPro; IPR008973; C2_GAB.
DR InterPro; IPR002420; PI3K_C2.
DR Pfam; PF00792; PI3K_C2; 1.
KW Kinase.
FT NON TER
SQ SEQUENCE 490 AA; 46593 MW; E60A14E45E9E4D48 CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 490;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 3 DGPTLR--EWISFC 14
DB 250 DGSTARWDEWLTFC 263

RESULT 36
Q8B1G9 PRELIMINARY; PRT; 509 AA.
AC Q8B1G9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30306P09 product:NA.K-ATPASE ALPHA 1 ISOFORM (EC 3.6.1.37) homolog.
DE Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EyeBall;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EyeBall;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EyeBall;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]

```

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Eyeball;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes";  
RL Genome Res. 10:1617-1630 (2000).  
[5]  
RN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-Format  
RT sequencing pipeline with 384 multiplexed capillary sequencer.";  
RL Genome Res. 10:1757-1771 (2000).  
[6]  
RN SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Eyeball;  
RX Adachi J., Aizawa K., Akiyama T., Aikawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuto M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirokane T.,  
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanukawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Niemi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Tota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK053751; BAC35507.1; -.  
DR HSSP; P06685; IM07.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0015662; F:ATPase activity; coupled to transmembrane m. . .; IEA.  
DR GO; GO:0016781; F:hydrolyase activity; IEA.  
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. . .; IEA.  
DR GO; GO:0006812; P:cation transport; IEA.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR InterPro; IPR006069; Cation ATPase.  
DR InterPro; IPR004014; Cation ATPase\_N.  
DR InterPro; IPR008250; E1-E2\_ATPase\_Reg.  
DR Pfam; PF00690; Cation\_ATPase\_N.1.  
DR Pfam; PF00122; E1-E2\_ATPase.1.  
DR PRINTS; PR00119; CATATPASE.  
DR PRINTS; PR00121; NAKATPASE.  
DR TIGRFAMs; TIGR01494; ATPase\_P-type; 2.  
DR PROSITE; PS00154; ATPASE\_E1-E2; UNKNOWN\_1.  
SQ SEQUENCE 509 AA; 55779 MW; 132C342CA9000E97 CRC64;  
Query Match 50.6%; Score 43; DB 2; Length 509;  
Best Local Similarity 60.0%; Pred. No. 93;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 5 PTLREWISFC 14  
DB 74 PTLREWVFC 83  
RESULT 37  
O80U28 PRELIMINARY; PRT; 960 AA.  
AC O80U28;  
DT 01-JUN-2003 (TREMblrel. 24, Created)  
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)

DE Atpla2 protein (Fragment).  
GN Name=Atpla2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CZECH II; TISSUE=Mammary tumor;  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marzula K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzyzanski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maitra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CZECH II; TISSUE=Mammary tumor;  
RA Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041774; AAH41774.1; -.  
DR HSSP; P06685; IM07.  
DR MGP; MGI:88106; Atpla2.  
DR GO; GO:0005890; C:sodium:potassium-exchanging ATPase complex; ISS.  
DR GO; GO:0005391; F:sodium:potassium-exchanging ATPase activity; ISS.  
DR GO; GO:0015991; P:ATP hydrolysis coupled proton transport; ISS.  
DR GO; GO:0030641; P:hydrogen ion homeostasis; ISS.  
DR GO; GO:0001504; P:neurotransmitter uptake; IMP.  
DR GO; GO:0006813; P:potassium ion transport; ISS.  
DR GO; GO:0006814; P:sodium ion transport; ISS.  
DR GO; GO:0030317; P:sperm motility; ISS.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR InterPro; IPR006069; Cation ATPase.  
DR InterPro; IPR004014; Cation ATPase\_N.  
DR InterPro; IPR008250; Denal\_like\_hydro.  
DR InterPro; IPR008250; E1-E2\_ATPase\_Reg.  
DR InterPro; IPR005775; Na/K\_ATPase\_alpha.  
DR Pfam; PF00689; Cation ATPase\_C.1.  
DR Pfam; PF00690; Cation ATPase\_C.1.  
DR Pfam; PF00122; E1-E2\_ATPase\_N.1.  
DR Pfam; PF00122; E1-E2\_ATPase.1.  
DR Pfam; PF00702; Hydrolyase.1.  
DR PRINTS; PR00119; CATATPASE.  
DR PRINTS; PR00121; NAKATPASE.  
DR TIGRFAMs; TIGR01406; ATPase-11C\_X-K; 1.  
DR TIGRFAMs; TIGR01494; ATPase\_P-type; 4.  
DR PROSITE; PS00154; ATPASE\_E1-E2; UNKNOWN\_1.  
FT NON\_TER 1  
SQ SEQUENCE 960 AA; 105641 MW; EA838C86819D0C45 CRC64;  
Query Match 50.6%; Score 43; DB 2; Length 960;  
Best Local Similarity 60.0%; Pred. No. 18+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 5 PTLREWISFC 14  
DB 22 PTLREWVFC 31

```

RESULT 38
ID Q91YV9 PRELIMINARY; PRT; 962 AA.
AC Q91YV9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013561; AAH13561.1; -.
DR HSSP; P06685; 1MO7.
DR GO; GO:0005890; C:sodium:potassium-exchanging ATPase complex; ISS.
DR GO; GO:0005391; P:sodium:potassium-exchanging ATPase activity; ISS.
DR GO; GO:0015991; P:ATP hydrolysis coupled proton transport; ISS.
DR GO; GO:0030641; P:hydrogen ion homeostasis; ISS.
DR GO; GO:0006813; P:potassium ion transport; ISS.
DR GO; GO:0006814; P:potassium ion transport; ISS.
DR GO; GO:0030317; P:sperm motility; ISS.
DR pfam; PF00689; Cation ATPase_C; 1.
DR pfam; PF00690; Cation ATPase_N; 1.
DR pfam; PF00122; E1-E2_ATPase; 1.
DR pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR TIGRFAWS; TIGR01494; ATPase_P-type; 1.
DR TIGRFAWS; TIGR01494; ATPase_P-type; 4.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 962 AA; 105826 MW; AB00A952F990AE45 CRC64;

```

```

Query Match 50.6%; Score 43; DB 2; Length 962;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 5 PTLREWISFC 14
Db 24 PTLPEWVKFC 33

```

```

RESULT 39
ID Q72419 PRELIMINARY; PRT; 1000 AA.
AC Q72419;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATPase Na+/K+ transporting alpha 4.
GN Name=ATP1A4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hlivko J.T., James P.F.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF506797; AAC07964.1; -.
DR HSSP; P06685; 1MO8.
DR GO; GO:0005890; C:sodium:potassium-exchanging ATPase complex; ISS.
DR GO; GO:0005391; P:sodium:potassium-exchanging ATPase activity; ISS.
DR GO; GO:0015991; P:ATP hydrolysis coupled proton transport; ISS.
DR GO; GO:0030641; P:hydrogen ion homeostasis; ISS.
DR GO; GO:0006813; P:potassium ion transport; ISS.
DR GO; GO:0006814; P:sodium ion transport; ISS.
DR GO; GO:0030317; P:sperm motility; ISS.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006068; Cation ATPase_C.
DR InterPro; IPR004014; Cation ATPase_N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR005775; Na/K_ATPase_alph.
DR pfam; PF00689; Cation ATPase_C; 1.
DR pfam; PF00690; Cation ATPase_N; 1.
DR pfam; PF00122; E1-E2_ATPase; 1.
DR pfam; PF00702; Hydrolyase; 1.
DR TIGRFAWS; TIGR01494; ATPase_P-type; 3.
DR TIGRFAWS; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
SQ SEQUENCE 1000 AA; 110943 MW; 5A9EBA0B24D482D1 CRC64;

```

```

Query Match 50.6%; Score 43; DB 2; Length 1000;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 5 PTLREWISFC 14
Db 63 PTLPEWVKFC 72

```

```

RESULT 40
ID Q98SL3 PRELIMINARY; PRT; 1009 AA.
AC Q98SL3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sodium/potassium pump alpha subunit.
OS Electrophorus electricus (Electric eel).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Gymnotiformes;
OC Electrophoridae; Electrophorus.
NCBI_TaxID=8005;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Electric organ;
RX MEDLINE=98068871; PubMed=9405797;
RA Kaya S., Yokoyama A., Imagawa T., Taniguchi K., Froehlich J.P.,
RA Albers R.W.;
RT "Cloning of the eel electroplex Na+,K(+)ATPase alpha subunit.";
RL Ann. N. Y. Acad. Sci. 834:129-131(1997).
RN [2]

```



Qy 5 PTLREWISFC 14  
Db 71 PTPPEWVKFC 80

## RESULT 42

AL3\_ORENO STANDARD; PRT; 1010 AA.

AC P58312; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Sodium/potassium-transferring ATPase alpha-3 chain (EC 3.6.3.9)

GN Name=ATP1A3; Oreochochromis mossambicus (Mozambique tilapia) (Tilapia mossambica).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Labroidae; Cichlidae; Oreochochromis.

OC NCBI\_Taxid=8127;

OX [1]

RP SEQUENCE FROM N.A.

RA Feng H.H., Leu J.H., Huang C.J., Hwang P.P.; Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: This is the catalytic component of the active enzyme, which catalyzes the hydrolysis of ATP coupled with the exchange of sodium and potassium ions across the plasma membrane. This action creates the electrochemical gradient of sodium and potassium ions, providing the energy for active transport of various nutrients.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+) (In) + K(+) (Out) = ADP + phosphate + Na(+) (Out) + K(+) (In).

CC -1- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and gamma.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily IIC.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AF109409; AAF75108.1; -

DR HSSP; P06685; IM07.

DR Interpro; IPR001757; ATPase\_E1-E2.

DR Interpro; IPR006069; Cation ATPase.

DR Interpro; IPR006068; Cation ATPase\_C.

DR Interpro; IPR004014; Cation ATPase\_N.

DR Interpro; IPR005834; Dehal\_like\_hydro.

DR Interpro; IPR008250; E1-E2\_ATPase\_reg.

DR Interpro; IPR005775; Na/K\_ATPase\_alph.

DR Pfam; PF00689; Cation\_ATPase\_C; 1.

DR Pfam; PF00122; E1-E2\_ATPase; 1.

DR Pfam; PF00702; Hydrolyase; 1.

DR PRINTS; PR00119; CATATPASE.

DR PRINTS; PR00121; NAKATPASE.

DR TIGRFAMs; TIGR01106; ATPase\_E1-E2; 1.

DR TIGRFAMs; TIGR01494; ATPase\_P-type; 5.

DR PROSITE; PS00154; ATPASE\_E1-E2; 1.

KM ATP-binding; Hydrolyase; Magnesium; Metal-binding; Multigene family; Cytoplasmic (Potential).

FT TRANSSEM 75 95 Potential.

FT TRANSSEM 96 118 Luminal (Potential).

FT TRANSSEM 119 139 Potential.

FT TRANSSEM 140 275 Cytoplasmic (Potential).

FT TRANSSEM 276 295 Potential.

FT TRANSSEM 296 307 Luminal (Potential).

FT TRANSSEM 308 325 Potential.

FT TRANSSEM 326 759 Cytoplasmic (Potential).

FT TRANSSEM 760 779 Potential (Potential).

FT TRANSSEM 780 789 Luminal (Potential).

FT TRANSSEM 790 810 Potential.

FT TRANSSEM 811 830 Cytoplasmic (Potential).

FT TRANSSEM 831 853 Potential.

FT TRANSSEM 854 905 Luminal (Potential).

FT TRANSSEM 906 925 Potential.

FT TRANSSEM 926 938 Cytoplasmic (Potential).

FT TRANSSEM 939 957 Potential.

FT TRANSSEM 958 972 Luminal (Potential).

FT TRANSSEM 973 993 Potential.

FT TRANSSEM 994 1010 Cytoplasmic (Potential).

FT TRANSSEM 10 133 Potential.

FT TRANSSEM 363 363 Poly-phs.

FT TRANSSEM 930 930 4-Aspartylphosphate intermediate (By similarity).

FT TRANSSEM 69 71 Phosphoserine (By PKA) (By similarity).

FT TRANSSEM 704 704 Binding of phosphoinositide-3 kinase (By similarity).

FT TRANSSEM 708 708 Magnesium (By similarity).

FT TRANSSEM 704 704 Magnesium (By similarity).

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;

FT TRANSSEM 1010 AA; 111506 MW; 9PEDB07C7547F0D1 CRC64;



```

DR Pfam: PF00122; E1-E2, ATPase; 1.
DR Pfam: PF00702; Hydrolyase; 1.
DR PRINTS; PRO0119; CATATPASE.
DR PRINTS; PRO0121; NAKATPASE.
DR TIGRFAMs: TIGR01106; ATPase-11C X-K; 1.
DR PROSITE: PS00154; ATPase_E1_E2; UNKNOWN_1.
DR PROSITE: PS00030; RRM_RNF_1; UNKNOWN_1.
SQ SEQUENCE 1012 AA; 111220 MW; FF93D65A24A06258 CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 2; Length 1012;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PTLRWISFC 14
DB 74 PTPPEWVKFC 83

RESULT 44
ID ALA3_HUMAN STANDARD; PRT; 1013 AA.
AC P13637; Q16732; Q16735; Q969K5;
DT 01-JAN-1990 (Rel. 13, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Sodium/potassium-transporting ATPase alpha-3 chain (EC 3.6.3.9)
DE (Sodium pump 3) (Na+/K+ ATPase 3) (Alpha1(III)).
GN Name=ATP1A3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8255304; PubMed=2838329; DOI=10.1016/0014-5793(88)81361-9;
RA Ovcchinikov Y.A., Monastyrskaya G.S., Brode N.E., Unshkaryov Y.A.,
RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Altkmetes R.L.,
RA Kostina M.B., Dulubova I.E., Kiyackin N.I., Grishin A.V.,
RA Modyanov N.N., Sverdlov E.D.;
RA "Family of human Na+, K+-ATPase genes. Structure of the gene for the
RA catalytic subunit (alpha III-form) and its relationship with
RA structural features of the protein.";
RT FEBS Lett. 233:87-94(1988).
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX PubMed=2834163;
RA Sverdlov E.D., Monastyrskaya G.S., Brode N.E., Unshkaryov Y.A.,
RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Altkmetes R.L.,
RA Kostina M.B., Dulubova I.E., Kiyackin N.I., Grishin A.V.,
RA Modyanov N.N., Ovcchinikov Y.A.;
RA "Family of human Na(+),K(+)-ATPase genes. Structure of the gene of
RA isoform alpha-III.";
RT Dokl. Akad. Nauk SSSR 297:1488-1494(1987).
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Magner L., Shemmen C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```

```

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Matria M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [4]
RN SEQUENCE OF 120-387; 494-538 AND 545-1013 FROM N.A.
RX MEDLINE=87162481; PubMed=3030810; DOI=10.1016/0014-5793(87)81467-9;
RA Ovcchinikov Y.A., Monastyrskaya G.S., Brode N.E., Altkmetes R.L.,
RA Unshkaryov Y.A., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
RA Dulubova I.E., Petrunkhin K.E., Gryshin A.V., Sverdlov E.E.,
RA Kiyackin N.I., Kostina M.B., Modyanov N.N., Sverdlov E.D.;
RA "The family of human Na+,K+-ATPase genes. A partial nucleotide
RA sequence related to the alpha-subunit.";
RT FEBS Lett. 213:73-80(1987).
RL [5]
RN ERRATUM.
RP Ovcchinikov Y.A., Monastyrskaya G.S., Brode N.E., Altkmetes R.L.,
RA Unshkaryov Y.A., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
RA Dulubova I.E., Petrunkhin K.E., Gryshin A.V., Sverdlov E.E.,
RA Kiyackin N.I., Kostina M.B., Modyanov N.N., Sverdlov E.D.;
RL FEBS Lett. 214:375-375(1987).
RN [6]
RP SEQUENCE OF 243-434 FROM N.A.
RX MEDLINE=87247232; PubMed=3035582; DOI=10.1016/0014-5793(87)80677-4;
RA Sverdlov E.D., Monastyrskaya G.S., Brode N.E., Unshkaryov Y.A.,
RA Altkmetes R.L., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
RA Dulubova I.E., Petrunkhin K.E., Gryshin A.V., Kiyackin N.I.,
RA Kostina M.B., Sverdlov E.E., Modyanov N.N., Ovcchinikov Y.A.;
RA "The family of human Na+,K+-ATPase genes. No less than five genes
RA and/or pseudogenes related to the alpha-subunit.";
RT FEBS Lett. 217:275-278(1987).
RL [7]
RN FUNCTION: This is the catalytic component of the active enzyme,
RN which catalyzes the hydrolysis of ATP coupled with the exchange of
RN sodium and potassium ions across the plasma membrane. This action
RN creates the electrochemical gradient of sodium and potassium ions,
RN providing the energy for active transport of various nutrients.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+)(In) + K(+)(Out) = ADP +
CC phosphate + Na(+)(Out) + K(+)(In).
CC -1- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and
CC gamma.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi-sib.ch).
CC -----
DR EMBL; M37457; AAA51798.1; -;
DR EMBL; M37436; AAA51798.1; JOINED.
DR EMBL; M37437; AAA51798.1; JOINED.
DR EMBL; M37438; AAA51798.1; JOINED.
DR EMBL; M37462; AAA51798.1; JOINED.
DR EMBL; M37440; AAA51798.1; JOINED.
DR EMBL; M37441; AAA51798.1; JOINED.
DR EMBL; M37442; AAA51798.1; JOINED.
DR EMBL; M37443; AAA51798.1; JOINED.
DR EMBL; M37444; AAA51798.1; JOINED.
DR EMBL; M37445; AAA51798.1; JOINED.
DR EMBL; M37447; AAA51798.1; JOINED.
DR EMBL; M37448; AAA51798.1; JOINED.
DR EMBL; M37449; AAA51798.1; JOINED.
DR EMBL; M37450; AAA51798.1; JOINED.
DR EMBL; M37451; AAA51798.1; JOINED.
DR EMBL; M37452; AAA51798.1; JOINED.
DR EMBL; M37453; AAA51798.1; JOINED.
DR EMBL; M37454; AAA51798.1; JOINED.

```



DR	EMBL	M37455	AAA51798.1	JOINED.	
DR	EMBL	M37456	AAA51798.1	JOINED.	
DR	EMBL	X12910	CAA31390.1	-	
DR	EMBL	X12911	CAA31390.1	JOINED.	
DR	EMBL	X12912	CAA31390.1	JOINED.	
DR	EMBL	X12913	CAA31390.1	JOINED.	
DR	EMBL	X12914	CAA31390.1	JOINED.	
DR	EMBL	X12915	CAA31390.1	JOINED.	
DR	EMBL	X12916	CAA31390.1	JOINED.	
DR	EMBL	X12917	CAA31390.1	JOINED.	
DR	EMBL	X12919	CAA31390.1	JOINED.	
DR	EMBL	X12920	CAA31390.1	JOINED.	
DR	EMBL	X12921	CAA31390.1	JOINED.	
DR	EMBL	X12922	CAA31390.1	JOINED.	
DR	EMBL	X12923	CAA31390.1	JOINED.	
DR	EMBL	M28286	AAA52285.1	-	
DR	EMBL	M28284	AAA52285.1	JOINED.	
DR	EMBL	M28285	AAA52285.1	JOINED.	
DR	EMBL	M28293	AAA52286.1	-	
DR	EMBL	M28287	AAA52286.1	JOINED.	
DR	EMBL	M35821	AAA52286.1	JOINED.	
DR	EMBL	M35822	AAA52286.1	JOINED.	
DR	EMBL	M28289	AAA52286.1	JOINED.	
DR	EMBL	M28290	AAA52286.1	JOINED.	
DR	EMBL	M28291	AAA52286.1	JOINED.	
DR	EMBL	M28292	AAA52286.1	JOINED.	
DR	EMBL	M27577	AAA58380.1	-	
DR	EMBL	M27570	AAA58380.1	JOINED.	
DR	EMBL	M27573	AAA58380.1	JOINED.	
DR	EMBL	BC009282	AAH09282.1	-	
DR	EMBL	BC009394	AAH09394.1	-	
DR	EMBL	BC015566	AAH15566.1	-	
DR	HSP	P06685	S00801		
DR	HSP	P06685	S00801		
DR	Genew	HGNC:801	ATP1A3		
DR	GO	GO:0005890	C:sodium/potassium-exchanging ATPase complex	ISS	
DR	GO	GO:0005351	P:sodium/potassium-exchanging ATPase activity	ISS	
DR	GO	GO:0010591	P:ATP hydrolysis coupled proton transport	ISS	
DR	GO	GO:0030641	P:hydrogen ion homeostasis	ISS	
DR	GO	GO:0006813	P:potassium ion transport	ISS	
DR	GO	GO:0006814	P:sodium ion transport	ISS	
DR	GO	GO:0030317	P:epERM motility	ISS	
DR	GO	GO:0006810	P:transport	TAS	
DR	InterPro	IPR001757	ATPase_E1-E2		
DR	InterPro	IPR006069	Cation ATPase		
DR	InterPro	IPR006068	Cation ATPase_C		
DR	InterPro	IPR004014	Cation ATPase_N		
DR	InterPro	IPR005334	DnaL-like_NydrO		
DR	InterPro	IPR008250	Na/E2 ATPase Reg		
DR	InterPro	IPR005775	Na/K ATPase alph		
DR	Pfam	PF00689	Cation ATPase_C	1	
DR	Pfam	PF00690	Cation ATPase_N	1	
DR	Pfam	PF00122	E1-E2 ATPase	1	
DR	Pfam	PF00702	HydroLase	1	
DR	PRINTS	PR00119	CATATPASE		
DR	PRINTS	PR00121	NAKATPASE		
DR	TIGRFAMs	TIGR01106	ATPase-IIIC_X-K	1	
DR	TIGRFAMs	TIGR01494	ATPase_P-type	5	
DR	PROSITE	PS00154	ATPase_E1-E2	1	
KW	ATP-binding	Hydrolase	Magnesium	Metal-binding	Multigene family
KW	Phosphorylation	Sodium/potassium transport	Transmembrane		
FT	DOMAIN	1	77	Cytoplasmic (Potential)	
FT	TRANSMEM	78	98	Potential	
FT	DOMAIN	99	121	luminal (Potential)	
FT	TRANSMEM	122	142	Potential	
FT	DOMAIN	143	278	Cytoplasmic (Potential)	
FT	TRANSMEM	279	298	Potential	
FT	DOMAIN	299	310	luminal (Potential)	
FT	TRANSMEM	311	328	Potential	
FT	DOMAIN	329	762	Cytoplasmic (Potential)	
FT	TRANSMEM	763	782	Potential	
FT	DOMAIN	783	792	luminal (Potential)	

FT	TRANSMEM	793	813	Potential.
Query Match		50.64;	Score 43;	DB 1; Length 1013;
Best Local Similarity		60.08;	Pred. No. 1.9e+02;	
Matches	6; Conservative	1;	Mismatches	3; Indels 0; Gaps 0;
Oy	5	PTLRWISFC 14		
Db	74	PTLRWISFC 83		
RESULT 45				
ALA3 MOUSE				
ID	ALA3 MOUSE	STANDARD;	PRT;	1013 AA.
AC	Q6P1C6;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Sodium/potassium-transporting ATPase alpha-3 chain (EC 3.6.3.9)			
DE	(Sodium pump 3) (Na+/K+ ATPase 3) (Alpha(III)).			
GN	Name=Atpl3;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RF	SEQUENCE FROM N.A.			
RC	TISSUE=Eye;			
RC	MEDLINE=22388257; PubMed=12477933; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loguélano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Boak S.A., McEwan P.O., McKernan K.O., Malek J.A., Gunaratne P.H.,			
RA	Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Faley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,			
RA	Schneerch A., Schein J.B., Jones S.J.W., Maier M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-1- FUNCTION: This is the catalytic component of the active enzyme,			
CC	which catalyzes the hydrolysis of ATP coupled with the exchange of			
CC	sodium and potassium ions across the plasma membrane. This action			
CC	creates the electrochemical gradient of sodium and potassium ions,			
CC	providing the energy for active transport of various nutrients (By			
CC	similarity).			
CC	-1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+) (In) + K(+) (Out) = ADP +			
CC	phosphate + Na(+) (Out) + K(+) (In).			
CC	-1- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and			
CC	gamma (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: Belongs to the cation transport ATPases family (P-type			
CC	ATPases). Subfamily IIC.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on how			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; BC034645; AAH34645.1; -			
DR	EMBL; BC037206; AAH37206.1; -			

```

DR EMBL; BC042894; AAA42894.1; -.
DR MGD; MGI:88107; AcpIa3.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Denal_Like_Hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR005775; Na/K_ATPase_alph.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; HydroIase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR TIGRFAMs; TIGR01106; ATPase-11C-X-K; 1.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
KM ATP-binding; Hydrolase; Magnesium; Metal-binding; Multigene family;
KW Phosphorylation; Sodium/potassium transport; Transmembrane.
FT DOMAIN 1 77 Cytoplasmic (Potential).
FT TRANSMEM 78 98 Potential.
FT DOMAIN 99 121 Lumenal (Potential).
FT TRANSMEM 122 142 Potential.
FT DOMAIN 143 228 Cytoplasmic (Potential).
FT TRANSMEM 229 279 Potential.
FT DOMAIN 279 310 Lumenal (Potential).
FT TRANSMEM 311 328 Potential.
FT DOMAIN 329 762 Cytoplasmic (Potential).
FT TRANSMEM 763 782 Potential.
FT DOMAIN 783 792 Lumenal (Potential).
FT TRANSMEM 793 813 Potential.
FT DOMAIN 814 833 Cytoplasmic (Potential).
FT TRANSMEM 834 856 Potential.
FT DOMAIN 857 908 Lumenal (Potential).
FT TRANSMEM 909 928 Potential.
FT DOMAIN 929 941 Cytoplasmic (Potential).
FT TRANSMEM 942 960 Potential.
FT DOMAIN 961 975 Lumenal (Potential).
FT TRANSMEM 976 996 Potential.
FT DOMAIN 997 1013 4-Aspartylphosphate intermediate (By
FT ACT_SITE 366 366 similarity).
FT MOD_RES 933 933 Phosphoserine (by PKA) (By similarity).
FT BINDING 72 74 Binding of phosphoinositide-3 kinase (By
FT METAL 707 707 similarity).
FT METAL 711 711 Magnesium (By similarity).
SQ SEQUENCE 1013 AA; 111690 MW; 72F051406284E8A8 CRC64;

Query Match 50.6%; Score 43; DB 1; Length 1013;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 5 PTLREWISFC 14
DB 74 PTPPEWVKFC 83

```

Search completed: September 1, 2005, 16:21:16  
 Job time : 54.0719 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:02 ; Search time 64.3597 Seconds  
(without alignments)  
84.131 Million cell updates/sec

Title: US-10-083-768-13

Perfect score: 73

Sequence: 1 IEGLPIRQMLARA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

## Database :

A\_Geneseq\_16Dec04:\*

1: geneseqp19808:\*

2: geneseqp19908:\*

3: geneseqp20008:\*

4: geneseqp20018:\*

5: geneseqp20028:\*

6: geneseqp20038:\*

7: geneseqp20038s:\*

8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	14	2	AAW09463 Thrombopo
2	73	100.0	14	2	AAW09468 Thrombopo
3	73	100.0	14	2	AAW3030 Thrombopo
4	73	100.0	14	2	AAW3034 Thrombopo
5	73	100.0	14	2	AAW36774 Thrombopo
6	73	100.0	14	2	AD124843 AF 12505
7	73	100.0	14	3	AA196515 Thrombopo
8	73	100.0	14	3	AA16962 TPO-mimet
9	73	100.0	14	4	AAU25827 Human thr
10	73	100.0	14	4	AAU26004 Human thr
11	73	100.0	14	5	ABR72853 TPO-mimet
12	73	100.0	14	5	ABP51669 Thrombopo
13	73	100.0	14	5	AAE18011 Human 11g
14	73	100.0	14	6	ABG71747 TPO recep
15	73	100.0	14	7	ABR62907 Thrombopo
16	73	100.0	14	7	ADC33697 Erythrope
17	73	100.0	14	7	ADN59652 Thrombopo
18	73	100.0	14	8	ADL27293 Amino aci
19	73	100.0	14	8	ADM72483 TPO mimet
20	73	100.0	14	8	ADM72483 Agonist T
21	73	100.0	15	2	AAW35416 Thrombopo
22	73	100.0	15	2	AAW36776 Thrombopo
23	73	100.0	15	3	AAW66712 Peptide c
24	73	100.0	15	3	AAW20684 Thrombocy
25	73	100.0	15	4	AAU25996 Human thr

26	73	100.0	15	4	AAU25831 Human thr
27	73	100.0	15	5	ABP51670 Thrombopo
28	73	100.0	15	7	ABR62908 Thrombopo
29	73	100.0	15	8	ADM72485 TPO mimet
30	73	100.0	15	8	ADM72479 TPO mimet
31	73	100.0	15	8	ADM72478 TPO mimet
32	73	100.0	15	8	ADM72533 TPO mimet
33	73	100.0	15	8	ADM72522 TPO mimet
34	73	100.0	15	8	ADM72523 TPO mimet
35	73	100.0	15	8	ADM72482 TPO mimet
36	73	100.0	15	8	ADQ16585 TPO mimet
37	73	100.0	16	2	AAW19534 Thrombopo
38	73	100.0	16	2	AAW3035 Thrombopo
39	73	100.0	16	2	AAW36771 Thrombopo
40	73	100.0	16	2	AAW36770 Peptide c
41	73	100.0	16	2	AAW66709 Peptide c
42	73	100.0	16	2	AAW66713 Peptide c
43	73	100.0	16	2	AAW66733 Peptide c
44	73	100.0	16	2	AAW66716 Peptide c
45	73	100.0	16	4	AAU26005 Human thr
46	73	100.0	16	4	AAU26043 Human thr
47	73	100.0	16	4	AAU25832 Human thr
48	73	100.0	16	8	ADM72532 TPO mimet
49	73	100.0	16	8	ADM72484 TPO mimet
50	73	100.0	18	3	AAW16957 PEGylated
51	73	100.0	18	3	AAW16956 PEGylated
52	73	100.0	18	5	ABP51687 TPO mimet
53	73	100.0	18	5	ABP51689 TPO mimet
54	73	100.0	18	5	ABP51688 TPO mimet
55	73	100.0	18	5	ABP51677 TPO mimet
56	73	100.0	18	5	ABP51686 TPO mimet
57	73	100.0	18	5	ABP51674 TPO mimet
58	73	100.0	18	5	ABP51683 TPO mimet
59	73	100.0	18	5	ABP51684 TPO mimet
60	73	100.0	18	5	ABP51683 TPO mimet
61	73	100.0	18	5	ABP51685 TPO mimet
62	73	100.0	18	5	ABP51691 TPO mimet
63	73	100.0	18	5	ABP51673 TPO mimet
64	73	100.0	18	5	ABP51690 TPO mimet
65	73	100.0	18	5	ABP51675 TPO mimet
66	73	100.0	18	5	ABP51692 TPO mimet
67	73	100.0	18	7	ADN59812 Thrombopo
68	73	100.0	18	8	ADQ16611 TPO mimet
69	73	100.0	18	8	ADQ16619 TPO mimet
70	73	100.0	18	8	ADQ16621 TPO mimet
71	73	100.0	18	8	ADQ16641 TPO mimet
72	73	100.0	18	8	ADQ16646 TPO mimet
73	73	100.0	18	8	ADQ16607 TPO mimet
74	73	100.0	18	8	ADQ16615 TPO mimet
75	73	100.0	18	8	ADQ16627 TPO mimet
76	73	100.0	18	8	ADQ16625 TPO mimet
77	73	100.0	18	8	ADQ16617 TPO mimet
78	73	100.0	18	8	ADQ16629 TPO mimet
79	73	100.0	18	8	ADQ16613 TPO mimet
80	73	100.0	18	8	ADQ16623 TPO mimet
81	73	100.0	18	8	ADQ16605 TPO mimet
82	73	100.0	18	8	ADQ16609 TPO mimet
83	73	100.0	19	5	ABR73391 TPO-mimet
84	73	100.0	19	5	ABR73390 FC-TWP pe
85	73	100.0	20	3	AAW18003 TPO-mimet
86	73	100.0	20	3	AAW17929 TPO-mimet
87	73	100.0	20	5	ABR73403 Thrombopo
88	73	100.0	21	7	ADN59819 TWP pepti
89	73	100.0	22	8	ADQ16714 Immunoglo
90	73	100.0	22	8	ADQ16713 Immunoglo
91	73	100.0	22	8	ADQ16709 Immunoglo
92	73	100.0	22	8	ADQ16706 Immunoglo
93	73	100.0	22	8	ADQ16699 TPO mimet
94	73	100.0	22	8	ADQ16712 Immunoglo
95	73	100.0	22	8	ADQ16707 Immunoglo
96	73	100.0	22	8	ADQ16711 Immunoglo
97	73	100.0	22	8	ADQ16708 Immunoglo
98	73	100.0	22	8	ADQ16708 Immunoglo

AAU25831	Human thr
ABP51670	Thrombopo
ABR62908	Thrombopo
ADM72485	TPO mimet
ADM72479	TPO mimet
ADM72478	TPO mimet
ADM72533	TPO mimet
ADM72522	TPO mimet
ADM72523	TPO mimet
ADM72482	TPO mimet
ADQ16585	TPO mimet
AAW19534	Thrombopo
AAW3035	Thrombopo
AAW36771	Thrombopo
AAW36770	Peptide c
AAW66713	Peptide c
AAW66733	Peptide c
AAW66716	Peptide c
AAU26005	Human thr
AAU26043	Human thr
AAU25832	Human thr
ADM72532	TPO mimet
ADM72484	TPO mimet
AAW16957	PEGylated
AAW16956	PEGylated
ABP51687	TPO mimet
ABP51689	TPO mimet
ABP51688	TPO mimet
ABP51677	TPO mimet
ABP51686	TPO mimet
ABP51674	TPO mimet
ABP51683	TPO mimet
ABP51684	TPO mimet
ABP51683	TPO mimet
ABP51685	TPO mimet
ABP51691	TPO mimet
ABP51673	TPO mimet
ABP51690	TPO mimet
ABP51675	TPO mimet
ABP51692	TPO mimet
ADN59812	Thrombopo
ADQ16611	TPO mimet
ADQ16619	TPO mimet
ADQ16621	TPO mimet
ADQ16641	TPO mimet
ADQ16646	TPO mimet
ADQ16607	TPO mimet
ADQ16615	TPO mimet
ADQ16627	TPO mimet
ADQ16625	TPO mimet
ADQ16617	TPO mimet
ADQ16629	TPO mimet
ADQ16613	TPO mimet
ADQ16623	TPO mimet
ADQ16605	TPO mimet
ADQ16609	TPO mimet
ABR73391	TPO-mimet
ABR73390	FC-TWP pe
AAW18003	TPO-mimet
AAW17929	TPO-mimet
ABR73403	Thrombopo
ADN59819	TWP pepti
ADQ16714	Immunoglo
ADQ16713	Immunoglo
ADQ16709	Immunoglo
ADQ16706	Immunoglo
ADQ16699	TPO mimet
ADQ16712	Immunoglo
ADQ16707	Immunoglo
ADQ16711	Immunoglo
ADQ16708	Immunoglo

99 73 100.0 22 8 ADQ16710  
100 73 100.0 28 3 AAb17285

## ALIGNMENTS

## RESULT 1

AAW09463  
ID AAW09463 standard; protein; 14 AA.

AC AAW09463;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;  
KM bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 1..14

/note= "Preferably linkages are selected from: -CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is lower alkyl"

FT Modified-site

/note= "Preferably N-terminus is selected from: -NR1; -NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide; benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3 substitutions on the phenyl ring selected from lower alkyl, lower alkoxy, chloro, bromo; where R and R1 are independently selected from hydrogen and lower alkyl"

FT Modified-site

14 /note= "Preferably C-terminus is -C(O)R2 where R2 is selected from hydroxy, lower alkoxy, and -NR3R4, where R3 and R4 are independently selected from hydrogen and lower alkyl, and where the nitrogen atom of the -NR3R4 group can optionally be the amine group of the N-terminus of the peptide forming a cyclic peptide"

PN W09640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

PS WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide  
PT mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.

XX Claim 18; Page 89; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)  
CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding  
CC affinity to TR as expressed by an IC50 of no more than about 100 nM. The  
CC compound (especially if modified, see features table) can be used for

CC treating patients suffering from haematological disorders and  
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
CC marrow transfusions. The peptide may also be used to maintain the  
CC proliferation and growth of TPO-dependent cell lines and for use in  
CC biological research, for detecting TPO receptors on living cells

XX Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1,4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLARA 14  
Db 1 IEPTLRQWLARA 14

## RESULT 2

AAW09468  
ID AAW09468 standard; protein; 14 AA.

AC AAW09468;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound peptide (part of a dimer).

XX Haematology; thrombocytopenia; TPO; TR; proliferation;  
KM bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

OS Synthetic.

Key Location/Qualifiers

FT Cross-links 14

/note= "Linked to the omega lys from AAW19534"

PN W09640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

PS WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide  
PT mimetic(s) - useful in treatment of haematological disorders, esp.  
PT thrombocytopenia resulting from chemotherapy, etc.

XX Claim 30; Page 91; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)  
CC receptor (TR). It is part of a dimer linked by the omega amino acid to  
CC the omega amino acid in the sequence in AAW19534. The compound can be  
CC used for treating patients suffering from haematological disorders and  
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
CC marrow transfusions. The peptide may also be used to maintain the  
CC proliferation and growth of TPO-dependent cell lines and for use in  
CC biological research, for detecting TPO receptors on living cells

XX Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAAARA 14  
 |||||  
 1 IEGPTLRQWLAAARA 14

RESULT 3  
 AAW33030  
 ID AAW33030 standard; peptide; 14 AA.

AC AAW33030;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN MO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PS Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a  
 CC molecular weight of less than 8000 Da and a TR binding affinity as  
 CC expressed by an IC50 of no more than about 100 microm. It can be used to  
 CC treat disorders which are susceptible to treatment with a thrombopoietin  
 CC agonist, preferably haematological disorders and thrombocytopenia  
 CC resulting from chemotherapy, radiation therapy or bone marrow  
 CC transfusions. It can also be used diagnostically, e.g. to investigate the  
 CC mechanism of thrombopoietin signal transduction and receptor activation,  
 CC or to maintain the proliferation and growth of thrombopoietin dependent  
 CC cell lines

SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAAARA 14  
 |||||  
 1 IEGPTLRQWLAAARA 14

RESULT 4  
 AAW33034  
 ID AAW33034 standard; peptide; 14 AA.

XX AAW33034;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN MO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX ) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PS Claim 30; Page 91; 106pp; English.  
 XX The present peptide binds the thrombopoietin receptor (TR), has a  
 CC molecular weight of less than 8000 Da and a TR binding affinity as  
 CC expressed by an IC50 of no more than about 100 microm. It can be used to  
 CC treat disorders which are susceptible to treatment with a thrombopoietin  
 CC agonist, preferably haematological disorders and thrombocytopenia  
 CC resulting from chemotherapy, radiation therapy or bone marrow  
 CC transfusions. It can also be used diagnostically, e.g. to investigate the  
 CC mechanism of thrombopoietin signal transduction and receptor activation,  
 CC or to maintain the proliferation and growth of thrombopoietin dependent  
 CC cell lines

SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAAARA 14  
 |||||  
 1 IEGPTLRQWLAAARA 14

RESULT 5  
 AAW36774  
 ID AAW36774 standard; peptide; 14 AA.

AC AAW36774;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KW haematological disorder; thrombocytopaenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 14  
 FT /note="NH2-Ala"  
 XX  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gares CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wighton NC;  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 XX  
 PS Example 9; Page 77; 106pp; English.  
 XX  
 CC The present peptide, which binds the thrombopoietin receptor (TP), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 XX  
 SQ Sequence 14 AA;  
 XX  
 QY  
 DB 1 IEPTLRQWLARA 14  
 1 IEPTLRQWLARA 14  
 1 IEPTLRQWLARA 14  
 XX  
 RESULT 6  
 AD124843  
 ID AD124843 standard; peptide; 14 AA.  
 XX  
 AC AD124843;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE AF 12505 as active moiety for pharmacologically active peptide.  
 KW pharmacologically active peptide conjugate; enzymatic cleavage; pain;  
 KW HIV; cancer; diabetes; incontinence; hypertension; amnesia;  
 KW Alzheimer's disease; fever; depression; sex hormone regulation;  
 KW eating disorder; schizophrenia; osteoporosis; insomnia;  
 KW Central nervous system disorder; contraceptive.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9946283-A1.  
 XX

PD 16-SEP-1999.  
 XX  
 PF 09-MAR-1999; 99WO-DK000118.  
 XX  
 PR 09-MAR-1998; 98DK-00000317.  
 XX  
 PA (ZEAL-) ZEALAND PHARM AS.  
 PI Larsen BD;  
 DR WPI; 1999-561659/47.  
 XX  
 PT New peptide conjugates used for treating, e.g. pain, HIV, depression,  
 PT schizophrenia, osteoporosis or insomnia.  
 XX  
 PS Claim 24; Page 90; 113pp; English.  
 XX  
 CC The invention relates to a novel pharmacologically active peptide  
 CC conjugate having a reduced tendency towards enzymatic cleavage comprises  
 CC X and Z, where: (a) X is a pharmacologically active peptide sequence; and  
 CC (b) Z is a stabilizing peptide sequence of 4-20 amino acid units  
 CC covalently bound to X, where each amino acid unit in the stabilizing  
 CC peptide sequence, Z being selected from Ala, Leu, Ser, Thr, Tyr, Asn,  
 CC Gln, Asp, Glu, Lys, Arg, His, Met, Orn, and amino acid units of formula -  
 CC NH-C(R1)(R2)-C(=O)- (1), where: R1 and R2 are H, 1-6C alkyl, phenyl, and  
 CC substituted selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfono, and carboxy, and phenyl and phenyl-methyl are optionally  
 CC substituted with 1-3 substituents selected from 1-6C-alkyl, 2-6C-alkenyl,  
 CC halogen, hydroxy, amino, cyano, nitro, sulfono, and carboxy, or R 1 and R  
 CC 2 together with the C atom to which they are bound form a cyclopentyl,  
 CC cyclohexyl or cycloheptyl ring, e.g. 2,4-diaminobutanoic acid and 2,3-  
 CC diaminopropanoic acid; the ratio between the half-life of the peptide  
 CC conjugate and the half-life of the corresponding active peptide sequence,  
 CC X, when treated with carboxypeptidase A or leucine aminopeptidase in  
 CC or in serum or plasma is at least about 2 (preferably at least about 10),  
 CC or when the pharmacologically active peptide X is not orally absorbed,  
 CC the conjugate is adsorbed, or a salt, with the proviso that the  
 CC pharmacologically active peptide conjugate is not selected from sequences  
 CC (AD124837)-(AD124841). The peptide conjugates can be used for treating  
 CC e.g. pain, HIV, cancer, diabetes, incontinence, hypertension, amnesia,  
 CC Alzheimer's disease, fever, depression, sex hormone regulation, eating  
 CC disorders, schizophrenia, osteoporosis or insomnia. They can also be used  
 CC for treating e.g. CNS disorders and as contraceptives. The conjugated  
 CC peptides are less susceptible to degradation by proteases compared to the  
 CC corresponding free pharmacologically active peptides. This sequence  
 CC represents a pharmacologically active peptide as the X part of the  
 CC peptide of the invention.  
 XX  
 SQ Sequence 14 AA;  
 XX  
 QY  
 DB 1 IEPTLRQWLARA 14  
 1 IEPTLRQWLARA 14  
 1 IEPTLRQWLARA 14  
 XX  
 RESULT 7  
 AAY96515  
 ID AAY96515 standard; peptide; 14 AA.  
 XX  
 AC AAY96515;  
 XX  
 DT 04-SEP-2000 (first entry)  
 XX  
 DE Thrombopoietin mimetic peptide.  
 KW Thrombopoietin; mimetic; TWP; platelet; megakaryocyte; production;  
 KW anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;  
 KW

KW cytotoxic/T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase; asclma;  
 KW thrombosis; pharmaceutical.  
 OS Synthetic.  
 PN WO200024782-A2.  
 XX  
 XX  
 XX PD 04-MAY-2000.  
 XX  
 XX PF 25-OCT-1999; 99MO-US025044.  
 XX  
 XX PR 23-OCT-1998; 98US-0105371P.  
 XX PR 22-OCT-1999; 99US-00428082.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Feige U, Liu C, Cheetham J, Boone TC;  
 XX WPI; 2000-350702/30.  
 DR  
 XX  
 PR Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.  
 XX  
 XX Claim 19, Page 189, 608pp; English.  
 XX  
 XX The present invention describes composition of matter (1) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (1) is:  
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)-d-P2, -(L1)-c-P1-  
 CC (L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytosolic, antiasthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AA65943 to AA65526 and AB16955 to  
 CC AA18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 XX Sequence 14 AA;  
 SQ  
 Query Match 100.0%; Score 73; DB 3; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Cx 1 IEGPTLRQWLARA 14  
 Db 1 IEGPTLRQWLARA 14  
 RESULT 9  
 AAU25827  
 ID AAU25827 standard; peptide: 14 AA.  
 XX  
 XX AAU25827;  
 AC  
 XX 17-DEC-2001 (first entry)  
 DT  
 XX Human thrombopoietin receptor (TPO-R) activator peptide #13.  
 DE  
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.  
 XX

OS	Homo sapiens.
XX	
FN	US6251864-B1.
PD	26-JUN-2001.
XX	
PF	01-MAR-2000; 2000US-00516704.
XX	
PR	07-JUN-1995; 95US-00478128.
PR	07-JUN-1995; 95US-00485301.
PR	07-JUN-1996; 96WO-US009623.
FR	15-AUG-1996; 96US-00699027.
XX	
PA	(GLAXO ) GLAXO GROUP LTD.
P1	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
P1	Balsubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;
P1	Yin Q;
DR	WPI; 2001-564142/63.
XX	
PT	Activating thrombopoietin receptors in cells, used to treat
PT	thrombocytopenia and hematological disorders, comprises contacting cells
PT	with peptides and peptide mimetics attached to hydrophilic polymers.
XX	
PS	Disclosure; Col 69-70; 128pp; English.
XX	
CC	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC	of activating thrombopoietin receptors in cells comprise contacting the
CC	cells with effective amounts of peptides and peptide mimetics attached to
CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC	as that due to chemotherapy, radiation therapy or bone-marrow
CC	transplantation and to prevent thrombocytopenia in patients at risk.The
CC	sequences are used to treat and prevent haematological disorders
CC	including thrombocytopenia and platelet disorders. They are used in vitro
CC	as unique tools for understanding the biological role of thrombopoietin
CC	(TPO) and to develop other compounds that bind to and activate the TPO
CC	receptor. The peptides can be used to detect TPO receptors on living
CC	cells and fixed cells, in biological fluids, in tissue homogenates, and
CC	in purified or natural biological materials. They may also be used for in
CC	situ staining, fluorescence-activated cell sorting, Western blotting, can
CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC	be used for in vitro expansion of megakaryocytes and their committed
CC	progenitors alone or in conjunction with additional cytokines
SQ	
XX	
SQ	Sequence 14 AA:
Query Match	100.0%; Score 73; DB 4; Length 14;
Best Local Similarity	100.0%; Pred. No. 1.4e-05;
Matches 14; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 IEGPTLRQWLARA 14 
Db	1 IEGPTLRQWLARA 14
RESULT 10	
AAU26004	
ID	AAU26004 standard; peptide; 14 AA.
XX	
AC	AAU26004;
XX	
DT	17-DEC-2001 (first entry)
XX	
DE	Human thrombopoietin receptor (TPO-R) activator peptide #190.
XX	
KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW	bone marrow transplantation; hematological disorder; platelet disorder;
KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

```

XX OS Homo sapiens.
XX PM US6251864-B1.
XX PD 26-JUN-2001.
XX PF 01-MAR-2000; 2000US-00516704.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00465101.
XX PR 07-JUN-1996; 96WO-US009623.
XX PR 15-AUG-1996; 96US-00699027.
XX PA (GLAXO ) GLAXO GROUP LTD.
XX PI Dower MJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;
XX PI Balesubramanian P, Wagerstrom CR, Hendren RM, Deprince RB, Podduturi S;
XX PI Yin Q;
XX DR WPI; 2001-564142/63.
XX PT Activating thrombopoietin receptors in cells, used to treat
XX PT thrombocytopenia and hematological disorders, comprises contacting cells
XX PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX PS Disclosure; Col 147; 128pp; English.
XX XX
XX XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX CC of activating thrombopoietin receptors in cells comprise contacting the
XX CC cells with effective amounts of peptides and peptide mimetics attached to
XX CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX CC as that due to chemotherapy, radiation therapy or bone-marrow
XX CC transplantation and to prevent thrombocytopenia in patients at risk.The
XX CC sequences are used to treat and prevent hematological disorders
XX CC including thrombocytopenia and platelet disorders. They are used in vitro
XX CC as unique tools for understanding the biological role of thrombopoietin
XX CC (TPO) and to develop other compounds that bind to and activate the TPO
XX CC receptor. The peptides can be used to detect TPO receptors on living
XX CC cells and fixed cells, in biological fluids, in tissue homogenates, and
XX CC in purified or natural biological materials. They may also be used for in
XX CC situ staining, fluorescence-activated cell sorting, Western blotting and
XX CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX CC be used for in vitro expansion of megakaryocytes and their committed
XX CC progenitors alone or in conjunction with additional cytokines
XX CC
XX SQ Sequence 14 AA;
XX
XX Query Match 100.0%; Score 73; DB 4; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1,4e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 1 IEPTIQLQWLARA 14
XX |||||
XX |||||
XX |||||
XX 1 IEPTIQLQWLARA 14
XX DB
XX
XX RESULT 11
XX ABB72853
XX ID ABB72853 standard; peptide; 14 AA.
XX AC ABB72853;
XX XX
XX DT 05-APR-2002 (first entry)
XX XX
XX DE TPO mimetic peptide SEQ ID NO:13.
XX XX
XX XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
XX KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
XX KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
XX KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
XX KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
XX KW

```



KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianemic; anorectic; antifertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200183525-A2.  
 PN 08-NOV-2001.  
 PD 02-MAY-2001; 2001WO-US014310.  
 PF 03-MAY-2000; 2000US-00563286.  
 PR (AMGE-) AMGEN INC.  
 PA Feige U, Liu C, Cheatham JC, Boone TC, Gudas JM;  
 PI WPI; 2002-130313/17.  
 DR Novel vehicle-peptide molecule or its multimers useful for treating  
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 XX diabetic retinopathy, obesity, sleep disorders and infertility.  
 PS Claim 39; Page 43; 176pp; English.  
 XX The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antirheumatic, antidiabetic, antidiabetic, ophthalmological,  
 CC antianemic, anorectic, antifertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
 CC mimetic compounds are useful for treating disorders characterised by low  
 CC red blood cell levels such as anaemia. The EPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 100.0%; Score 73; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAA 14  
 Db 1 IEGPTLRQWLAA 14  
 RESULT 12  
 ABP51669 ID ABP51669 standard; peptide; 14 AA.  
 XX  
 AC ABP51669;  
 XX  
 DT 01-OCT-2002 (first entry)  
 XX

DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:1.  
 XX TPO, EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200246238-A2.  
 PN 13-JUN-2002.  
 PD 05-DEC-2001; 2001WO-US047656.  
 PF 05-DEC-2000; 2000US-0251448P.  
 PR 04-MAY-2001; 2001US-0288889P.  
 PR 29-MAY-2001; 2001US-0294068P.  
 XX (ALEX-) ALEXION PHARM INC.  
 PA Bowdish KS, Barbas-Frederickson S, Renshaw M;  
 PI WPI; 2002-566610/50.  
 DR A novel immunogen molecule comprising a region in which amino acid  
 XX residues corresponding to at least a portion of the complementary  
 PT determining region are replaced or fused with an erythropoietin or  
 PT thrombopoietin mimetic.  
 PS Claim 16; Page 6; 113pp; English.  
 XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to  
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABO73288 to ABO73377 and ABP51669 to ABP51696 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 SQ Sequence 14 AA;  
 XX  
 Query Match 100.0%; Score 73; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAA 14  
 Db 1 IEGPTLRQWLAA 14  
 RESULT 13  
 AAE18011 ID AAE18011 standard; peptide; 14 AA.  
 XX  
 AC AAE18011;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX



XX OS Synthetic.  
 XX PN WO2003059251-A2.  
 XX PD 24-JUL-2003.  
 XX PF 22-OCT-2002; 2002WO-US033991.  
 XX PR 22-OCT-2001; 2001US-0344614P.  
 XX PR 19-SRP-2002; 2002US-0412455P.  
 XX PA (SCRI ) SCRIPPS RES INST.  
 XX PI Barbas CF, Rader C, Sinha SC, Lerner R;  
 XX DR WPI; 2003-636673/60.  
 XX PT Antibody targeting compound useful e.g. for diagnostic immunoassays and  
 XX PT treating microbial diseases comprises targeting or biological agent  
 XX PT covalently linked to combining site of the antibody.  
 XX PS Example 7; Page 62; 56pp; English.  
 XX CC The present sequence is that of thrombopoietin (TPO) mimetic peptide  
 CC AF12305, which mimics the activity of recombinant TPO. The invention  
 CC provides antibody targeting compounds that are used to reprogram the  
 CC specificity of an antibody. The antibody targeting compound is linked to  
 CC the combining site of the antibody, such that the modified antibody takes  
 CC on the binding specificity of the targeting agent. In an example from the  
 CC invention, a TPO receptor targeting antibody compound was prepared by  
 CC covalently linking peptide AF12305 to aldolase monoclonal antibody 38C2.  
 CC The TPO receptor targeting antibody compound can be used to treat  
 CC thrombocytopaenia resulting from chemotherapy and bone marrow  
 CC transplantation  
 XX SQ Sequence 14 AA;  
 XX QY  
 XX DB 1 IEGPTLRQWLAAARA 14  
 1 IEGPTLRQWLAAARA 14  
 Query Match 100.0%; Score 73; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAAARA 14  
 DB 1 IEGPTLRQWLAAARA 14  
 RESULT 16  
 ID ADC33697 standard; peptide; 14 AA.  
 XX AC ADC33697;  
 XX DT 18-DEC-2003 (first entry)  
 XX DE Erythropoietin receptor/erythropoietin consensus peptide SEQ ID NO:1.  
 XX KM chimeric retrovirus envelope protein; ecotropic envelope protein;  
 XX KM cytoskeletal; gene therapy; cancer.  
 XX OS Synthetic.  
 XX OS WO2003076596-A2.  
 XX PN 18-SEP-2003.  
 XX PD 07-MAR-2003; 2003WO-US007323.  
 XX PR 08-MAR-2002; 2002US-0362655P.  
 XX PA (UYMA-) UNIV MASSACHUSETTS.  
 XX PI Green MR, Gollan TV;

XX DR WPI; 2003-722332/68.  
 XX DR New chimeric retrovirus envelope protein comprising an ecotropic envelope  
 XX PT protein and a heterologous short peptide ligand inserted within the  
 XX PT ecotropic envelope protein useful for treating cancer.  
 XX PS Disclosure; SEQ ID NO 1; 42pp; English.  
 XX CC The present invention describes a chimeric retrovirus envelope protein  
 CC (1) comprising an ecotropic envelope protein and a heterologous short  
 CC peptide ligand inserted within the ecotropic envelope protein. Also  
 CC described: (1) a nucleic acid molecule comprising a sequence encoding the  
 CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic  
 CC acid sequence encoding the chimeric envelope protein; (3) a recombinant  
 CC retroviral particle comprising a chimeric envelope protein comprising a  
 CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)  
 CC identifying a nucleic acid sequence encoding the chimeric envelope  
 CC protein that alters viral tropism; (5) delivering a nucleic acid sequence  
 CC to a cell; and (6) treating cancer. (1) has cytostatic activity and can  
 CC be used in gene therapy. The chimeric retrovirus envelope protein is  
 CC useful for treating cancer, which comprises providing a cancer cell, e.g.  
 CC human cancer cell and infecting the cancer cell with a virus, e.g.  
 CC retrovirus comprising the chimeric envelope protein comprising a  
 CC heterologous short peptide ligand and a therapeutically useful gene, e.g.  
 CC encoding thymidine kinase. The present sequence represents an  
 CC erythropoietin receptor/erythropoietin consensus peptide, which is given  
 CC in the exemplification of the present invention.  
 XX SQ Sequence 14 AA;  
 XX QY  
 XX DB 1 IEGPTLRQWLAAARA 14  
 1 IEGPTLRQWLAAARA 14  
 Query Match 100.0%; Score 73; DB 7; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAAARA 14  
 DB 1 IEGPTLRQWLAAARA 14  
 RESULT 17  
 ID ADNS9652 standard; peptide; 14 AA.  
 XX AC ADNS9652;  
 XX DT 01-JUL-2004 (first entry)  
 XX DE Thrombopoietin mimetic peptide (TMP), seq id 1.  
 XX KM Haemostatic; antihaemic; immunosuppressive; platelet;  
 XX KM transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;  
 XX KM TMP; c-mpl receptor; platelet precursor; megakaryocyte;  
 XX KM thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;  
 XX KM autoimmune haemolytic anaemia; Hughes' s syndrome;  
 XX KM lupoid thrombocytopaenia.  
 XX OS Homo sapiens.  
 XX OS WO2003031589-A2.  
 XX PN 17-APR-2003.  
 XX PD 11-OCT-2002; 2002WO-US032552.  
 XX PR 11-OCT-2001; 2001US-0328666P.  
 XX PR 10-OCT-2002; 2002US-00269806.  
 XX PA (AMGR-) AMGEN INC.  
 XX PA Min H, Sitney KC, Hartley C;  
 XX PI WPI; 2003-403101/38.

XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and  
PT which stimulate the production of platelets and/or the production of  
PT platelet precursors, useful for treating thrombocytopenia.  
XX  
25 Disclosure; SEQ ID NO 1; 126pp; English.

The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that binds to the c-mpl (mpl) receptor, and which stimulates the production of platelets and/or the production of platelet precursors, is new. Further disclosed is a composition of matter (II) that binds to an mpl receptor and a pharmaceutical composition comprising (II) and a carrier. The pharmaceutical composition of the invention is useful for treating thrombocytopenia in an animal, and for increasing megakaryocytes or platelets in a patient. The TMP of the invention is useful for treating conditions involving a megakaryocyte and/or platelet deficiency, e.g. disease conditions involving thrombocytopenia such as aplastic anaemia, autoimmune thrombocytopenia, drug induced immune thrombocytopenia, autoimmune haemolytic anaemia, Hughes' syndrome and lupoid thrombocytopenia. The TMP of the invention is also useful for maintaining the viability or storage life of platelets and/or megakaryocytes and its derived cells. The compounds demonstrate an improved ability to bind to and/or trigger transmembrane signal through, i.e. activating, the mpl receptor the compounds have superior thrombopoietic activity, i.e. the ability to stimulate, in vivo and in vitro, the production of platelets and/or megakaryocytic activity i.e. the ability to stimulate, in vivo and in vitro, the production of platelet precursors. Further, certain of the compounds also exhibit superior therapeutic properties, such as improved plasma half-life, biological activity and in vivo circulation time. The current sequence represents a TMP of the invention.

Query Match	100.0%	Score 73;	DB 7;	Length 14;
Best Local Similarity	100.0%;	Pred. No.	1.4e-05;	
Matches 14; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

```
QY      1 IEGPTLRQWLAARA 14
         |||||
Db      1 IEGPTLRQWLAARA 14
```

## RESULT 18

ID ADL27293 standard; protein; 14 AA.

AC ADL27293;

DT 03-JUN-2004 (first entry)

DE Amino acid sequence of a thrombopoietin agonist peptide.

KW fusion protein; C4bp; alpha chain; systemic lupus erythematosus.

## OS Homo sapiens.

OS Synthetic.

PN WO2004020639-A2.

PD 11-MAR-2004

PF 12-AUG-2003; 2003WO-EP008928.

PR 14-AUG-2002; 2002EP-00292043.

PA (AVID-) AVIDIS SA.

PI Garnier L, Hill F, Julien M;

DR WPI; 2004-239202/22.

PT Obtaining a recombinant fusion protein, useful for treating lupus,

PT comprises providing a prokaryotic host cell carrying a nucleic acid  
PT encoding the recombinant protein operably linked to a promoter functional  
PT in the prokaryotic cell.

PS Claim 8; Page 48; 69pp; English.

CC The specification describes a method for obtaining a recombinant fusion  
CC protein comprising a scaffold of a C-terminal core protein of C4bp alpha  
CC chain, where the recombinant fusion protein is capable of forming  
CC multimers in soluble form in a prokaryotic host cell. The method  
CC comprises providing a prokaryotic host cell carrying a nucleic acid  
CC encoding the recombinant protein operably linked to a promoter functional  
CC in the prokaryotic cell, culturing the host cell under conditions where  
CC the recombinant protein is expressed, and recovering the recombinant  
CC protein where the protein is recovered in multimeric form without  
CC performing a scaffold refolding step. The protein is useful for treating  
CC systemic lupus erythematosus. The present sequence represents a  
CC thrombopoietin agonist peptide, which is used to produce fusion proteins  
CC of the invention.

Query Match	100.0%	Score 73;	DB 8;	length 14;
Best Local Similarity	100.0%	Pred. No. 1.4e-05;		
Matches 14; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0.

```
QY 1 IEGPTLRQWLAARA 14
    |||||
Db 1 IEGPTLRQWLAARA 14
```

## RESULT 19

ID ADM72483 standard; peptide; 14 AA.

AC ADM72483 ;

DT 17-JUN-2004 (first entry)

**TPO mimetic peptide fragment.**

KM stem cell therapy; HSC; transplantation; engraftment; mimetic. TPO; haematopoietic stem cell; thrombopoietin; haemostatic;

Synthetic.

PN WO2004026332-A1.

PD 01-APR-2004

PF 18-SEP-2003; 2003WO-US029701

PR 18-SEP-2002; 2002US-0411700P

PR 18-SEP-2002; 2002US-0411779P.

PA (THRE-) 3-DIMENSIONAL PHARM INC.

PI Kaushansky K, Macdonald BR;

DR WPI; 2004-283153/26

PT Increasing hematopoietic stem cell production in subject, useful in  
PT reducing the incidence of delayed primary engraftment, comprises  
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a  
PT subject.

PS Disclosure; Fig 2; 32pp; English

CC The invention relates to a method (M1) for increasing haematopoietic stem  
CC cell production in a subject which involves administering a  
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
CC another method (M2) of providing haematopoietic stem cells to a subject  
CC which involves administering a TPO mimetic compound to a subject to

CC enhance expansion of a stem cell population within bone marrow and/or  
 CC mobilize stem cells in peripheral circulation, harvesting one or more of  
 CC the bone marrow stem cells or the stem cells in the peripheral  
 CC circulation, and transplanting the harvested stem cells into the subject.  
 CC A method (M3) is also provided for reducing a time to engraftment  
 CC following reinfusion of stem cells in a subject, involves administering a  
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem  
 CC cell population within bone marrow and/or mobilizing the stem cells in  
 CC peripheral circulation, harvesting one or more of the bone marrow stem  
 CC cells or one or more of the stem cells in the peripheral circulation, and  
 CC transplanting the one or more harvested stem cells into the subject. TPO  
 CC mimetic compounds are disclosed as peptides, including cyclic or modified  
 CC peptides. (M1) is useful for increasing haematopoietic stem cell  
 CC production in a subject e.g., human. (M3) is useful for reducing time to  
 CC engraftment following reinfusion of stem cells, reducing the incidence of  
 CC delayed primary engraftment, reducing the incidence of secondary failure  
 CC of platelet production and reducing the time of platelet and/or  
 CC neutrophil engraftment following reinfusion of stem cells in a subject.  
 CC (M1) is also useful for increasing the number of stem cells from a donor  
 CC whose cells are then used for rescue of recipient subject. Also useful in  
 CC the treatment of thrombocytopenia. (M1) enables transplantation to  
 CC proceed in patients who would not otherwise be considered as candidates  
 CC because of unacceptably high risk of failed engraftment, reduces the  
 CC number of aphereses required to generate a minimum acceptable harvest,  
 CC reduces the incidence of primary and secondary failure of engraftment by  
 CC increasing the number of haematopoietic stem cells (HSCs) available for  
 CC transplantation and reduces the time required for primary engraftment.  
 CC The present sequence represents an example of TPO mimetic peptide  
 CC fragment.

XX Sequence 14 AA;

Query Match 100.0%; Score 73; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLARA 14  
 |||||  
 1 IEGPTLRQWLARA 14

RESULT 20

ADQ16584 ID ADQ16584 standard; peptide; 14 AA.

XX AC ADQ16584;

XX DT 09-SEP-2004 (first entry)

DE Agonist TPO mimetic peptide SEQ ID NO:1.

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KM immunotherapy; thrombocytopenia.

OS Unidentified.

XX WO2004050017-A2.

PD 17-JUN-2004.

XX 17-NOV-2003; 2003WO-US036894.

PR 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M;

XX WPI; 2004-460973/43.

PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic

PT or a TPO mimetic, useful for treating thrombocytopenia.

XX Claim 8; SEQ ID NO 1; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide.

XX Sequence 14 AA;

Query Match 100.0%; Score 73; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLARA 14  
 |||||  
 1 IEGPTLRQWLARA 14

RESULT 21

AAW35416 ID AAW35416 standard; peptide; 15 AA.

XX AC AAW35416;

XX DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KM haematological disorder; thrombocytopenia; chemotherapy;  
 KM radiation therapy; bone marrow transfusion; diagnosis;  
 KM signal transduction; receptor activation; cell culture.

OS Synthetic.

XX Key Location/Qualifiers

FT Cross-links 1 /note="linked via disulfide bond to Cys1 of identical

FT Modified-site 15 /note="NH2-Ala"

XX WO9640750-A1.

XX 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO ) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Caira SE, Duffin DJ, Gates CM, Johnson SS;

XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 73; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines  
 CC  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 73; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAAARA 14  
 |||||  
 DB 2 IEGPTLRQWLAAARA 15  
 RESULT 22  
 AAM36776  
 ID AAM36776 standard; peptide; 15 AA.  
 XX  
 AC AAM36776;  
 XX  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Thrombopoietin receptor binding peptide.  
 XX  
 KM Thrombopoietin receptor; binding peptide; treatment; agonist;  
 KM haematological disorder; thrombocytopenia; chemotherapy;  
 KW radiation therapy; bone marrow transfusion; diagnosis;  
 KW signal transduction; receptor activation; cell culture.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Cross-links 1  
 FT /note="linked via disulfide bond to Cys1 of identical  
 FT peptide"  
 FT Modified-site 15  
 FT /note="NH2-Ala"  
 FT  
 FT  
 PN WO9640750-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PR 07-JUN-1996; 96WO-US009623.  
 XX  
 PR 07-JUN-1995; 95US-00478128.  
 PR 07-JUN-1995; 95US-00485301.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;  
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;  
 PI  
 DR WPI; 1997-052226/05.  
 XX  
 PT Peptides and peptide mimetics which bind to and activate the  
 PT thrombopoietin receptor - useful in treatment of haematological  
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.  
 PS  
 PS Example 9; Page 77; 106pp; English.  
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be  
 CC used to treat disorders which are susceptible to treatment with a  
 CC thrombopoietin agonist, preferably haematological disorders and  
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone  
 CC marrow transfusions. It can also be used diagnostically, e.g. to  
 CC investigate the mechanism of thrombopoietin signal transduction and  
 CC receptor activation, or to maintain the proliferation and growth of  
 CC thrombopoietin dependent cell lines

XX  
 SQ Sequence 15 AA;  
 Query Match 100.0%; Score 73; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAAARA 14  
 |||||  
 DB 2 IEGPTLRQWLAAARA 15  
 RESULT 23  
 AAM66712  
 ID AAM66712 standard; peptide; 15 AA.  
 XX  
 AC AAM66712;  
 XX  
 DT 01-DEC-1998 (first entry)  
 XX  
 DE Peptide chain of compound which binds to the thrombopoietin receptor.  
 XX  
 KM thrombopoietin receptor; haematological disorder; screening; agonist;  
 KW assay; megakaryocyte; blood disorder; thrombocytopenia; TPO.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..14  
 FT /note="thrombopoietin receptor agonist"  
 FT Modified-site 15  
 FT /note="Epsilon amino group of Lys, in its amide form, is  
 FT attached to another peptide chain identical to the region  
 FT (residues 1 to 14) of this peptide"  
 FT  
 FT  
 PN WO9825965-A2.  
 XX  
 PD 18-JUN-1998.  
 XX  
 PR 09-DEC-1997; 97WO-EP006850.  
 XX  
 PR 11-DEC-1996; 96US-00764640.  
 XX  
 PA (GLAXO ) GLAXO GROUP LTD.  
 XX  
 PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
 PI Balasubramanian P, Wagstrom CR, Hendren RM, Depirince RB, Podduturi S;  
 PI Yin Q;  
 XX  
 DR WPI; 1998-377261/32.  
 XX  
 PT New peptide compound(s) which can bind and activate thrombopoietin  
 PT receptor - may be used in treating haematological disorders and in  
 PT methods for screening for new thrombopoietin receptor agonists.  
 PS  
 PS Claim 2; Page 60; 78pp; English.  
 CC The invention relates to peptide compounds composed of two peptide chains  
 CC attached to each of the amino groups of a single lys in the amide form.  
 CC The compounds are of formula (Pept1)(Pept2)K(NH2), where Pept1 is of  
 CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9-X10; and Pept2 is of  
 CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9'-X10'. X1 = H or acyl; X2  
 CC = Gly or Sar (sarcosine); X3 = Arg, Ala, Nle (norleucine) or N-  
 CC acetyllysine; X4 = Gln or Glu; X5 = Trp, L-1-naphthylalanine or Phe; X6 =  
 CC Ala, 5-aminopentanoic acid or 2-aminobutyric acid; X7 = Ala,  
 CC diphenylalanine, or is absent; X8 = Arg, p- amino-phenylalanine, N-  
 CC acetyl-lysine, or is absent; X9, X9' = Ala, beta Ala, N-methyl-alanine,  
 CC Sar, or is absent; X10, X10' = beta Ala or is absent. The new peptides  
 CC are capable of binding to, and activating, the thrombopoietin (TPO)  
 CC receptor. They may be used in vitro as tools for understanding the  
 CC biological role of TPO. They may be used as competitive binders in assays  
 CC to screen for new TPO receptor agonists. They may be used as reagents for  
 CC detecting TPO receptors in living cells, biological fluids, etc. They may

CC be used to maintain growth and proliferation of TPO-dependent cells and  
 CC for in vitro expansion of megakaryocytes. They may be used to activate  
 CC TPO receptors in vivo, e.g., to treat blood disorders or  
 CC thrombocytopenia associated with bone marrow transplants, radiotherapy  
 CC or chemotherapy. The present sequence represents a specific example of  
 CC (Pep)K(NH<sub>2</sub>)

XX  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGLTIRQWLAA 14  
 |||||  
 1 IEGLTIRQWLAA 14

Db

RESULT 24  
 AAB20684 standard; peptide; 15 AA.

XX  
 AC AAB20684;

DT 20-DEC-2000 (first entry)

XX Thrombocyte generation promoting peptide.

XX Thrombocyte; promotion; generation.

XX Unidentified.

XX Key Location/Qualifiers

FT Modified-site 15 /note="optionally amidated; optionally attached to the C  
 FT -terminal cysteine of a similar peptide"

PN CNI254718-A.

XX 31-MAY-2000.

XX 20-NOV-1998; 98CN-00125011.

XX 20-NOV-1998; 98CN-00125011.

XX (BIOL-) INST BIOLOGICAL ENG CHINESE ACAD MILITAR.

PI Cheng D, Li C, Huang P;

XX WPI; 2000-533568/49.

PT Active peptide.

PS Claim 1; Page 1; Spp; Chinese.

XX The present invention discloses an active peptide which promotes  
 CC thrombocyte generation. The active peptide can be synthesised by a  
 CC polypeptide solid-phase synthesis method, and has the monomer sequence of  
 CC IEGLTIRQWLAAAC and the amidated peptide chain structure of  
 CC IEGLTIRQWLAAAC-NH<sub>2</sub>. Its activity is increased by 20 times for its  
 CC monomer, or by 10 times for the amidated peptide chain compared with the  
 CC monomer, or by 100 times for its dimer compared with its monomer

XX Sequence 15 AA;

Query Match 100.0%; Score 73; DB 3; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGLTIRQWLAA 14  
 |||||  
 1 IEGLTIRQWLAA 14

Db 1 IEGLTIRQWLAA 14

RESULT 25

AAU25996 standard; peptide; 15 AA.

XX AAU25996;

DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #182.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
 KW bone marrow transplantation; haematological disorder; platelet disorder;  
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
 XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

XX US6251864-B1.

PD 26-JUN-2001.

PF 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAXO) GLAXO GROUP LTD.

XX Power WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprience RB, Poddaturi S,

PI Yin Q;

XX WPI; 2001-564142/63.

XX Disclosure; Col 143-144; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
 CC of activating thrombopoietin receptors in cells comprise contacting the  
 CC cells with effective amounts of peptides and peptide mimetics attached to  
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
 CC as that due to chemotherapy, radiation therapy or bone-marrow  
 CC transplantation and to prevent thrombocytopenia in patients at risk. The  
 CC sequences are used to treat and prevent haematological disorders  
 CC including thrombocytopenia and platelet disorders. They are used in vitro  
 CC as unique tools for understanding the biological role of thrombopoietin  
 CC (TPO) and to develop other compounds that bind to and activate the TPO  
 CC receptor. The peptides can be used to detect TPO receptors on living  
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
 CC in purified or natural biological materials. They may also be used for in  
 CC situ staining, fluorescence-activated cell sorting, Western blotting and  
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
 CC be used for in vitro expansion of megakaryocytes and their committed  
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 15 AA;

Query Match 100.0%; Score 73; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGLTIRQWLAA 14  
 |||||  
 2 IEGLTIRQWLAA 15

Db 2 IEGLTIRQWLAA 15



RESULT 26  
AAU25831  
ID AAU25831 standard; peptide; 15 AA.  
XX  
AC AAU25831;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human thrombopoietin receptor (TPO-R) activator peptide #17.  
XX  
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;  
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
KW bone marrow transplantation; haematological disorder; platelet disorder;  
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.  
XX  
OS Homo sapiens.  
XX  
PN US6251864-B1.  
XX  
PD 26-JUN-2001.  
XX  
PF 01-MAR-2000; 2000US-00516704.  
XX  
PR 07-JUN-1995; 95US-00478128.  
XX 07-JUN-1995; 95US-00485301.  
PR 07-JUN-1996; 96WO-US009623.  
PR 15-AUG-1996; 96US-00699027.  
XX  
PA (GLAXO) GLAXO GROUP LTD.  
XX  
PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ,  
PI Balasubramanian P, Wagerstrom CR, Hendren RM, Depirnce RB, Poddurti S;  
PI Yin Q;  
XX  
DR WPI; 2001-564142/63.  
XX  
PT Activating thrombopoietin receptors in cells, used to treat  
PT thrombocytopenia and hematological disorders, comprises contacting cells  
PT with peptides and peptide mimetics attached to hydrophilic polymers.  
XX  
PS Claim 1; Col 69-70; 128pp; English.  
XX  
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that  
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods  
CC of activating thrombopoietin receptors in cells comprise contacting the  
CC cells with effective amounts of peptides and peptide mimetics attached to  
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such  
CC as that due to chemotherapy, radiation therapy or bone-marrow  
CC transplantation and to prevent thrombocytopenia in patients at risk. The  
CC sequences are used to treat and prevent haematological disorders  
CC including thrombocytopenia and platelet disorders. They are used in vitro  
CC as unique tools for understanding the biological role of thrombopoietin  
CC (TPO) and to develop other compounds that bind to and activate the TPO  
CC receptor. The peptides can be used to detect TPO receptors on living  
CC cells and fixed cells, in biological fluids, in tissue homogenates, and  
CC in purified or natural biological materials. They may also be used for in  
CC situ staining, fluorescence-activated cell sorting, Western blotting and  
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can  
CC be used for in vitro expansion of megakaryocytes and their committed  
CC progenitors alone or in conjunction with additional cytokines  
XX  
SQ Sequence 15 AA:  
XX  
Query Match 100.0%; Score 73; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 IEPTLRQWLAARA 14  
RESULT 27  
ABP51670  
ID ABP51670 standard; peptide; 15 AA.  
XX  
AC ABP51670;  
XX  
DT 01-OCT-2002 (first entry)  
XX  
DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:2.  
XX  
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
KW complementarity determining region; immunoglobulin; antianemic;  
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.  
XX  
OS Homo sapiens.  
XX  
PN WO200246238-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 05-DEC-2001; 2001WO-US047656.  
XX  
PR 05-DEC-2000; 2000US-0251448P.  
XX 04-MAY-2001; 2001US-0288889P.  
PR 29-MAY-2001; 2001US-0294068P.  
XX  
PA (ALEXON) PHARM INC.  
XX  
PI Bowdish KS, Barbados-Frederickson S, Renshaw M;  
XX  
DR WPI; 2002-566610/60.  
XX  
PT A novel immunogen molecule comprising a region in which amino acid  
PT residues corresponding to at least a portion of the complementary  
PT determining region are replaced or fused with an erythropoietin or  
PT thrombopoietin mimetic.  
XX  
PS Claim 19; Page 6; 113pp; English.  
XX  
CC The present invention describes an immunoglobulin molecule or its fragment  
CC (I) comprising a region where amino acid residues corresponding to at  
CC least a portion of the complementary determining region (CDR) are  
CC replaced or fused with biologically active peptides e.g. a peptide  
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
CC that is flanked with proline at its carboxy terminus. (I) has  
CC antianemic, haemostatic and nephrotropic activities, and can be used as  
CC a stimulator of proliferation, differentiation and maturation of  
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
CC for stimulating proliferation, differentiation or growth of  
CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
CC promegakaryocytes or megakaryocytes, which results in increased platelet  
CC production. (I) with a region where amino acid residues corresponding to  
CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
CC production of red blood cells, where (I) is contacted with haematopoietic  
CC stem cells or their progenitors. (I) is useful for diagnostics or  
CC therapeutics, in cell isolation strategies, and for treating patients  
CC suffering from deficiency in cell populations caused by disease,  
CC disorders or treatments related to the suppression of haematopoiesis.  
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in  
CC the exemplification of the present invention  
XX  
SQ Sequence 15 AA:  
XX  
Query Match 100.0%; Score 73; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



DB 1 IEGPTLRQWLAAARA 14

RESULT 28  
ABR62908  
ID ABR62908 standard; peptide; 15 AA.

AC ABR62908;

DT 04-DEC-2003 (first entry)

DE Thrombopoietin mimetic peptide AFI2505.

KM Thrombopoietin; mimetic; thrombocytopenia; antibody targeting.

OS Synthetic.

PN WO2003059251-A2.

PD 24-JUL-2003.

PP 22-OCT-2002; 2002WO-US033991.

PR 22-OCT-2001; 2001US-0344614P.

PR 19-SEP-2002; 2002US-0412455P.

PA (SCRI ) SCRIAPS RES INST.

PI Barbas CF, Rader C, Sinha SC, Lerner R;

PT MPI; 2003-636673/60.

PT Antibody targeting compound useful e.g. for diagnostic immunoassays and

PT creating microbial diseases comprises targeting or biological agent

PT covalently linked to combining site of the antibody.

PS Example 7; Page 62; 56pp; English.

XX The present sequence is that of thrombopoietin (TPO) mimetic peptide

CC AFI2505, modified to include an N-terminal Cys residue. AFI2505 mimics

CC the activity of recombinant TPO. The invention provides antibody

CC targeting compounds that are used to reprogram the specificity of an

CC antibody. The antibody targeting compound is linked to the combining

CC site of the antibody, such that the modified antibody takes on the binding

CC specificity of the targeting agent. In an example from the invention, a

CC TPO receptor targeting antibody compound was prepared by covalently

CC linking Cys-modified peptide AFI2505 to aldolase monoclonal antibody 38C2

CC using a maleimide-diketone linker. The resulting TPO receptor targeting

CC antibody compound can be used to treat thrombocytopenia resulting from

CC chemotherapy and bone marrow transplantation

CC Sequence 15 AA;

QY Query Match 100.0%; Score 73; DB 7; Length 15;

DB Best Local Similarity 100.0%; Pred. No. 1.5e-05; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IEGPTLRQWLAAARA 14

2 IEGPTLRQWLAAARA 15

RESULT 29

ADM72485

ID ADM72485 standard; peptide; 15 AA.

AC ADM72485;

DT 17-JUN-2004 (first entry)

DE TPO mimetic peptide fragment.

KM TPO; haematopoietic stem cell; thrombopoietin; haemostatic;

KM stem cell therapy; HSC; transplantation; engraftment; mimetic.

OS Synthetic.

FT Key

FT Modified-site

FT 15

FT /label= bAla

FT /note= "beta-alanine"

FT WO2004026332-A1.

PN 01-APR-2004.

PP 18-SEP-2003; 2003WO-US029701.

PR 18-SEP-2002; 2002US-0411700P.

PR 18-SEP-2002; 2002US-0411779P.

PA (THRE-) 3-DIMENSIONAL PHARM INC.

PI Kausansky K, Macdonald BR;

PT MPI; 2004-283153/26.

PT Increasing hematopoietic stem cell production in subject, useful in

PT reducing the incidence of delayed primary engraftment, comprises

PT administering a Thrombopoietin mimetic compound e.g., a peptide to a

PT subject.

PS Disclosure; Fig 2; 32pp; English.

XX The invention relates to a method (M1) for increasing haematopoietic stem

CC cell production in a subject which involves administering a

CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is

CC another method (M2) of providing haematopoietic stem cells to a subject

CC which involves administering a TPO mimetic compound to a subject to

CC enhance expansion of a stem cell population within bone marrow and/or

CC mobilize stem cells in peripheral circulation, harvesting one or more of

CC the bone marrow stem cells or the stem cells in the peripheral

CC circulation, and transplanting the harvested stem cells into the subject.

CC A method (M3) is also provided for reducing a time to engraftment

CC following reinfusion of stem cells in a subject, involves administering a

CC TPO mimetic compound to the subject, enhancing the expansion of the stem

CC cell population within bone marrow and/or mobilizing the stem cells in

CC peripheral circulation, harvesting one or more of the bone marrow stem

CC cells or one or more of the stem cells in the peripheral circulation, and

CC transplanting the one or more harvested stem cells into the subject. TPO

CC mimetic compounds are disclosed as peptides, including cyclic or modified

CC peptides. (M1) is useful for increasing haematopoietic stem cell

CC production in a subject e.g., human. (M3) is useful for reducing time to

CC engraftment following reinfusion of stem cells, reducing the incidence of

CC delayed primary engraftment, reducing the incidence of secondary failure

CC of platelet production and reducing the time of platelet and/or

CC neutrophil engraftment following reinfusion of stem cells in a subject.

CC (M1) is also useful for increasing the number of stem cells from a donor

CC whose cells are then used for rescue of recipient subject. Also useful in

CC the treatment of thrombocytopenia. (M1) enables transplantation to

CC proceed in patients who would not otherwise be considered as candidates

CC because of unacceptably high risk of failed engraftment, reduces the

CC number of adherences required to generate a minimum acceptable harvest,

CC reduces the incidence of primary and secondary failure of engraftment by

CC increasing the number of haematopoietic stem cells (HSCs) available for

CC transplantation and reduces the time required for primary engraftment.

CC The present sequence represents an example of TPO mimetic peptide

CC fragment.

XX Sequence 15 AA;

QY Query Match 100.0%; Score 73; DB 8; Length 15;

DB Best Local Similarity 100.0%; Pred. No. 1.5e-05; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IEGPTLRQWLAAARA 14

2 IEGPTLRQWLAAARA 15

QY 1 IEGPTLRQWLARA 14  
 |||||  
 Db 1 IEGPTLRQWLARA 14

RESULT 30  
 ADM72479  
 ID ADM72479 standard; peptide, 15 AA.  
 AC ADM72479;  
 XX  
 DT 17-JUN-2004 (first entry)  
 DE TPO mimetic peptide fragment.  
 XX  
 XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;  
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note= "beta-alanine"  
 FT  
 PN WO2004026332-A1.  
 XX  
 PD 01-APR-2004.  
 PD  
 PF 18-SEP-2003; 2003WO-US029701.  
 XX  
 PF 18-SEP-2002; 2002US-0411700P.  
 PR 18-SEP-2002; 2002US-0411779P.  
 XX  
 XX (THRE-) 3-DIMENSIONAL PHARM INC.  
 XX  
 PI Kausanský K, Macdonald BR;  
 XX  
 DR WPI; 2004-283153/26.  
 XX  
 PT Increasing hematopoietic stem cell production in subject, useful in  
 PT reducing the incidence of delayed primary engraftment, comprises  
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a  
 PT subject.  
 XX  
 PS Disclosure; Fig 2; 32pp; English.  
 XX  
 XX The invention relates to a method (M1) for increasing haematopoietic stem  
 CC cell production in a subject which involves administering a  
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
 CC another method (M2) of providing haematopoietic stem cells to a subject  
 CC which involves administering a TPO mimetic compound to a subject to  
 CC enhance expansion of a stem cell population within bone marrow and/or  
 CC mobilize stem cells in peripheral circulation, harvesting one or more of  
 CC the bone marrow stem cells or the stem cells in the peripheral  
 CC circulation, and transplanting the harvested stem cells into the subject.  
 CC A method (M3) is also provided for reducing a time to engraftment  
 CC following reinfusion of stem cells in a subject, involves administering a  
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem  
 CC cell population within bone marrow and/or mobilizing the stem cells in  
 CC peripheral circulation, harvesting one or more of the bone marrow stem  
 CC cells or one or more of the stem cells in the peripheral circulation, and  
 CC transplanting the one or more harvested stem cells into the subject. TPO  
 CC mimetic compounds are disclosed as peptides, including cyclic or modified  
 CC peptides. (M1) is useful for increasing haematopoietic stem cell  
 CC production in a subject e.g., human. (M3) is useful for reducing time to  
 CC engraftment following reinfusion of stem cells, reducing the incidence of  
 CC delayed primary engraftment, reducing the time of platelet and/or  
 CC neutrophil production and reducing the time of platelet and/or  
 CC neutrophil engraftment following reinfusion of stem cells in a subject.  
 CC (M1) is also useful for increasing the number of stem cells from a donor  
 CC whose cells are then used for rescue of recipient subject. Also useful in  
 CC the treatment of thrombocytopenia. (M1) enables transplantation to  
 CC proceed in patients who would not otherwise be considered as candidates

CC because of unacceptably high risk of failed engraftment, reduces the  
 CC number of aphereses required to generate a minimum acceptable harvest.  
 CC increases the incidence of primary and secondary failure of engraftment by  
 CC increasing the number of haematopoietic stem cells (HSCs) available for  
 CC transplantation and reduces the time required for primary engraftment.  
 CC The present sequence represents an example of TPO mimetic peptide  
 CC fragment.  
 XX  
 SQ Sequence 15 AA;  
 XX  
 XX

Query Match 100.0%; Score 73; DB 8; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLARA 14  
 |||||  
 Db 1 IEGPTLRQWLARA 14

RESULT 31  
 ADM72478  
 ID ADM72478 standard; peptide, 15 AA.  
 XX  
 AC ADM72478;  
 XX  
 DT 17-JUN-2004 (first entry)  
 DE TPO mimetic peptide fragment.  
 XX  
 DE TPO; haematopoietic stem cell; thrombopoietin; haemostatic;  
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note= "Iys (15) is linked to one copy of the TPO mimetic  
 FT peptide through the alpha amino group and to a second  
 FT copy of the peptide (not shown) via the omega amino  
 FT group"  
 FT  
 PN WO2004026332-A1.  
 XX  
 PD 01-APR-2004.  
 PD  
 PF 18-SEP-2003; 2003WO-US029701.  
 XX  
 PF 18-SEP-2002; 2002US-0411700P.  
 PR 18-SEP-2002; 2002US-0411779P.  
 XX  
 XX (THRE-) 3-DIMENSIONAL PHARM INC.  
 XX  
 PI Kausanský K, Macdonald BR;  
 XX  
 DR WPI; 2004-283153/26.  
 XX  
 PT Increasing hematopoietic stem cell production in subject, useful in  
 PT reducing the incidence of delayed primary engraftment, comprises  
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a  
 PT subject.  
 XX  
 PS Disclosure; Fig 2; 32pp; English.  
 XX  
 XX The invention relates to a method (M1) for increasing haematopoietic stem  
 CC cell production in a subject which involves administering a  
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
 CC another method (M2) of providing haematopoietic stem cells to a subject  
 CC which involves administering a TPO mimetic compound to a subject to  
 CC enhance expansion of a stem cell population within bone marrow and/or  
 CC mobilize stem cells in peripheral circulation, harvesting one or more of  
 CC the bone marrow stem cells or the stem cells in the peripheral  
 CC circulation, and transplanting the harvested stem cells into the subject.  
 CC A method (M3) is also provided for reducing a time to engraftment



WO2004026532-A1.  
 01-APR-2004.  
 18-SEP-2003; 2003WO-US029701.  
 18-SEP-2002; 2002US-0411700P.  
 18-SEP-2002; 2002US-0411779P.  
 (THRE-) 3-DIMENSIONAL PHARM INC.  
 Kaushansky K, Macdonald BR;  
 WPI; 2004-283153/26.  
 Increasing hematopoietic stem cell production in subject, useful in  
 reducing the incidence of delayed primary engraftment, comprises  
 administering a Thrombopoietin mimetic compound e.g., a peptide to a  
 subject.  
 Disclosure; Fig 2, 32pp; English.  
 The invention relates to a method (M1) for increasing haematopoietic stem  
 cell production in a subject which involves administering a  
 Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
 another method (M2) of providing haematopoietic stem cells to a subject  
 which involves administering a TPO mimetic compound to a subject to  
 enhance expansion of a stem cell population within bone marrow and/or  
 mobilize stem cells in peripheral circulation, harvesting one or more of  
 the bone marrow stem cells or the stem cells in the peripheral  
 circulation, and transplanting the harvested stem cells into the subject.  
 A method (M3) is also provided for reducing a time to engraftment  
 following reinfusion of stem cells in a subject, involves administering a  
 TPO mimetic compound to the subject, enhancing the expansion of the stem  
 cell population within bone marrow and/or mobilizing the stem cells in  
 peripheral circulation, harvesting one or more of the bone marrow stem  
 cells or one or more of the stem cells in the peripheral circulation, and  
 transplanting the one or more harvested stem cells into the subject. TPO  
 mimetic compounds are disclosed as peptides, including cyclic or modified  
 peptides. (M1) is useful for increasing haematopoietic stem cell  
 production in a subject e.g., human. (M3) is useful for reducing time to  
 engraftment following reinfusion of stem cells, reducing the incidence of  
 delayed primary engraftment, reducing the incidence of secondary failure  
 of platelet production and reducing the time of platelet and/or  
 neutrophil engraftment following reinfusion of stem cells in a subject.  
 (M1) is also useful for increasing the number of stem cells from a donor  
 whose cells are then used for rescue of recipient subject. Also useful in  
 the treatment of thrombocytopenia. (M1) enables transplantation to  
 proceed in patients who would not otherwise be considered as candidates  
 because of unacceptably high risk of failed engraftment, reduces the  
 number of aphereses required to generate a minimum acceptable harvest,  
 reduces the incidence of primary and secondary failure of engraftment by  
 increasing the number of hematopoietic stem cells (HSCs) available for  
 transplantation and reduces the time required for primary engraftment.  
 The present sequence represents an example of TPO mimetic peptide  
 fragment.  
 Sequence 15 AA;  
 Query Match 100.0%; Score 73; DB 8; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 1 IEGPLTROWLARA 14  
 1 IEGPLTROWLARA 14  
 RESULT 34  
 ADM72523  
 ID ADM72523 standard; peptide; 15 AA.  
 AC ADM72523;

17-JUN-2004 (first entry)  
TPO mimetic peptide fragment.  
TPO; haematopoietic stem cell; thrombopoietin; haemostatic;  
stem cell therapy; HSC; transplantation; engraftment; mimetic.  
Synthetic.  
Key Location/Qualifiers  
Modified-site 15 /note= "beta-alanine"  
WO200402632-A1.  
01-APR-2004.  
18-SEP-2003; 2003WO-US029701.  
18-SEP-2002; 2002US-0411700P.  
18-SEP-2002; 2002US-0411799P.  
(THRE-) 3-DIMENSIONAL PHARM INC.  
Kaushansky K, Macdonald BR;  
WPI; 2004-283153/26.  
Increasing hematopoietic stem cell production in subject, useful in  
reducing the incidence of delayed primary engraftment, comprises  
administering a Thrombopoietin mimetic compound e.g., a peptide to a  
subject.  
Disclosure; Fig 2; 32pp; English.  
The invention relates to a method (M1) for increasing haematopoietic stem  
cell production in a subject which involves administering a  
Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
another method (M2) of providing haematopoietic stem cells to a subject  
which involves administering a TPO mimetic compound to a subject to  
enable expansion of a stem cell population within bone marrow and/or  
mobilize stem cells in peripheral circulation, harvesting one or more of  
the bone marrow stem cells or the stem cells in the peripheral  
circulation, and transplanting the harvested stem cells into the subject.  
A method (M3) is also provided for reducing a time to engraftment  
following reinfusion of stem cells in a subject, involves administering a  
TPO mimetic compound to the subject, enhancing the expansion of the stem  
cell population within bone marrow and/or mobilizing the stem cells in  
peripheral circulation, harvesting one or more of the bone marrow stem  
cells or one or more of the stem cells in the peripheral circulation, and  
transplanting the one or more harvested stem cells into the subject. TPO  
mimetic compounds are disclosed as peptides, including cyclic or modified  
peptides. (M1) is useful for increasing haematopoietic stem cell  
production in a subject e.g., human. (M3) is useful for reducing time to  
engraftment following reinfusion of stem cells, reducing the incidence of  
delayed primary engraftment, reducing the incidence of secondary failure  
of platelet production and, reducing the time of platelet and/or  
neutrophil engraftment following reinfusion of stem cells in a subject.  
(M1) is also useful for increasing the number of stem cells from a donor  
whose cells are then used for rescue of recipient subject. Also useful in  
the treatment of thrombocytopenia. (M1) enables transplantation to  
proceed in patients who would not otherwise be considered as candidates  
because of unacceptably high risk of failed engraftment, reduces the  
number of aphereses required to generate a minimum acceptable harvest,  
reduces the incidence of primary and secondary failure of engraftment by  
increasing the number of haematopoietic stem cells (HSCs) available for  
transplantation and reduces the time required for primary engraftment.  
The present sequence represents an example of TPO mimetic peptide  
fragment.  
Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
 |||||  
 1 IEGPTLRQWLAAARA 14

DB 1 IEGPTLRQWLAAARA 14

RESULT 35  
 ADM72482  
 ID ADM72482 standard; peptide; 15 AA.  
 XX  
 AC ADM72482;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE TPO mimetic peptide fragment.  
 XX  
 KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;  
 KM stem cell therapy; HSC; transplantation; engraftment; mimetic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note="lys (15) is linked to one copy of the TPO mimetic  
 FT peptide through the alpha amino group and to a second  
 FT copy of the peptide (not shown) via the omega amino  
 FT group"  
 XX  
 PN WO2004026332-A1.  
 XX  
 PD 01-APR-2004.  
 XX  
 PF 18-SEP-2003; 2003WO-US029701.  
 XX  
 PR 18-SEP-2002; 2002US-0411700P.  
 PR 18-SEP-2002; 2002US-0411779P.  
 XX  
 PA (THRE-) 3-DIMENSIONAL PHARM INC.  
 PI Kaushansky K, Macdonald BR;  
 PI WPI; 2004-283153/26.  
 XX  
 DR WPI; 2004-283153/26.  
 XX  
 PT Increasing hematopoietic stem cell production in subject, useful in  
 PT reducing the incidence of delayed primary engraftment, comprises  
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a  
 PT subject.  
 XX  
 PS Disclosure; Fig 2; 32pp; English.  
 XX  
 CC The invention relates to a method (M1) for increasing haematopoietic stem  
 CC cell production in a subject which involves administering a  
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is  
 CC another method (M2) of providing haematopoietic stem cells to a subject  
 CC which involves administering a TPO mimetic compound to a subject to  
 CC enhance expansion of a stem cell population within bone marrow and/or  
 CC mobilize stem cells in peripheral circulation, harvesting one or more of  
 CC the bone marrow stem cells or the stem cells in the peripheral  
 CC circulation, and transplanting the harvested stem cells into the subject.  
 CC A method (M3) is also provided for reducing a time to engraftment  
 CC following reinfusion of stem cells in a subject, involves administering a  
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem  
 CC cell population within bone marrow and/or mobilizing the stem cells in  
 CC peripheral circulation, harvesting one or more of the bone marrow stem  
 CC cells or one or more of the stem cells in the peripheral circulation, and  
 CC transplanting the one or more harvested stem cells into the subject. TPO  
 CC mimetic compounds are disclosed as peptides, including cyclic or modified  
 CC peptides. (M1) is useful for increasing haematopoietic stem cell  
 CC production in a subject e.g., human. (M3) is useful for reducing time to  
 CC engraftment following reinfusion of stem cells, reducing the incidence of

CC delayed primary engraftment, reducing the incidence of secondary failure  
 CC of platelet production and reducing the time of platelet and/or  
 CC neutrophil engraftment following reinfusion of stem cells in a subject.  
 CC (M1) is also useful for increasing the number of stem cells from a donor  
 CC whose cells are then used for rescue of recipient subject. Also useful in  
 CC the treatment of thrombocytopenia. (M1) enables transplantation to  
 CC proceed in patients who would not otherwise be considered as candidates  
 CC because of unacceptably high risk of failed engraftment, reduces the  
 CC number of aphereses required to generate a minimum acceptable harvest,  
 CC increases the incidence of primary and secondary failure of engraftment by  
 CC transplanting the number of haematopoietic stem cells (HSCs) available for  
 CC The present sequence represents an example of TPO mimetic peptide  
 CC fragment.

Sequence 15 AA;  
 XX  
 SQ

Query Match 100.0%; Score 73; DB 8; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
 |||||  
 1 IEGPTLRQWLAAARA 14

DB 1 IEGPTLRQWLAAARA 14

RESULT 36  
 ADQ16585  
 ID ADQ16585 standard; peptide; 15 AA.  
 XX  
 AC ADQ16585;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE TPO mimetic peptide SEQ ID NO:2.  
 XX  
 DE Immunoglobulin; complementarity determining region; CDR; peptide mimetic;  
 KM erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;  
 KM immunotherapy; thrombocytopenia.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004050017-A2.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PF 17-NOV-2003; 2003WO-US036894.  
 XX  
 PR 02-DEC-2002; 2002US-00307724.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 PI Bowdish KS, Frederickson S, Renshaw M;  
 PI WPI; 2004-460973/43.  
 XX  
 DR WPI; 2004-460973/43.  
 XX  
 PT New immunoglobulin molecule comprising a region, where two  
 PT complementarity determining regions (CDRs) are replaced with EPO mimetic  
 PT or a TPO mimetic, useful for treating thrombocytopenia.  
 XX  
 PS Disclosure; SEQ ID NO 2; 107pp; English.  
 XX  
 CC The invention relates to a novel immunoglobulin molecule or its fragment  
 CC comprising a region where amino acid residues corresponding to at least a  
 CC portion of a two complementarity determining regions (CDRs) are replaced  
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and  
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the  
 CC invention has immunosuppressive activity, and may have a use in  
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or  
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow  
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.  
 CC The present sequence represents a TPO mimetic peptide.

SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLARA 14  
 |||||  
 DB 1 IEPTLRQWLARA 14

RESULT 37

AAW19534  
 ID AAW19534 standard; protein; 16 AA.

AAW19534;

10-SEP-1997 (first entry)

Thrombopoietin receptor binding compound peptide (part of a dimer).

Haematology; thrombocytopenia; TPO; TR; proliferation;

bone marrow transfusion; chemotherapy; radiation therapy.

Synthetic.

Key Location/Qualifiers

Modified-site 15 /label= bala

Cross-links 16 /note= "linked to the omega Ala in AAW09468"

Modified-site 16 /note= "In amide form"

WO9640189-A1.

19-DEC-1996.

05-JUN-1996; 96WO-US008998.

07-JUN-1995; 95US-00472371.

07-JUN-1995; 95US-00473604.

07-JUN-1995; 95US-00476168.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00484090.

07-JUN-1995; 95US-00485301.

(GLAX ) GLAXO GROUP LTD.

Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Mathaeakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

Claim 30; Page 91; 106pp; English.

The present sequence is a compound which binds to thrombopoietin (TPO) receptor (TR). It is part of a dimer linked by the omega amino acid to the omega amino acid in the sequence in AAW09468. The compound can be used for treating patients suffering from haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. The peptide may also be used to maintain the proliferation and growth of TPO-dependent cell lines and for use in biological research, for detecting TPO receptors on living cells

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLARA 14  
 |||||  
 DB 1 IEPTLRQWLARA 14

RESULT 38

AAW33035  
 ID AAW33035 standard; peptide; 16 AA.

AAW33035;

11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;

haematological disorder; thrombocytopenia; chemotherapy;

radiation therapy; bone marrow transfusion; diagnosis;

signal transduction; receptor activation; cell culture.

Synthetic.

Key Location/Qualifiers

Cross-links 14 /note= "epsilon amino group of Lys16 linked to terminal carboxy group of AAW33034"

Modified-site 15 /label= bala

WO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAX ) GLAXO GROUP LTD.

Dower WJ, Barrett RM, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;

Mathaeakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

Claim 30; Page 91; 106pp; English.

Peptides and peptide mimetics which bind to and activate the thrombopoietin receptor - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

The present peptide binds the thrombopoietin receptor (TR), has a molecular weight of less than 8000 Da and a TR binding affinity as expressed by an IC50 of no more than about 100 microm. It can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 39
AAW36775
ID AAW36775 standard; peptide; 16 AA.
AC
XX AAW36775;
DT
XX 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1. .16
XX Modified-site 16
XX /note= "NH2-Cys"
XX
XX PA WO9640750-A1.
XX PN 19-DEC-1996.
XX PD
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX
XX PT Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX PS Example 9; Page 77; 106pp; English.
XX
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX CC investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX CC thrombopoietin dependent cell lines
XX
SQ Sequence 16 AA;

```

```

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 IEPTLRQWLAAARA 14
   |||||
Db 2 IEPTLRQWLAAARA 15

```

```

RESULT 40
AAW36771
ID AAW36771 standard; peptide; 16 AA.
AC
XX AAW36771;
XX
DT 11-MAR-1998 (first entry)

```

```

XX
DE Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1. .16
XX Modified-site 16
XX /note= "NH2-Cys"
XX
XX PA WO9640750-A1.
XX PN 19-DEC-1996.
XX PD
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX PI Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX
XX PT Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX PS Example 9; Page 76; 106pp; English.
XX
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX CC thrombopoietin agonist, preferably haematological disorders and
XX CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX CC investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX CC thrombopoietin dependent cell lines
XX
SQ Sequence 16 AA;

```

```

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 IEPTLRQWLAAARA 14
   |||||
Db 2 IEPTLRQWLAAARA 15

```

```

RESULT 41
AAW66709
ID AAW66709 standard; peptide; 16 AA.
AC
XX AAW66709;
XX
DT 01-DEC-1998 (first entry)
XX
DE Peptide chain of compound attached to hydrophilic polymer.
XX
XX Thrombopoietin receptor; haematological disorder; screening; agonist;
XX assay; megakaryocyte; blood disorder; thrombocytopenia; IPO.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers

```









CC (Pep1)K(INH2)  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAAARA 14  
| | | | | | | | | |  
| | | | | | | | | |  
Db 1 IEGPTLRQWLAAARA 14

RESULT 45  
AAU26005  
ID AAU26005 standard; peptide; 16 AA.  
XX  
AC AAU26005;

DT 17-DEC-2001 (first entry)  
XX

DE Human thrombopoietin receptor (TPO-R) activator peptide #191.  
XX

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;  
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;  
KW bone marrow transplantation; haematological disorder; platelet disorder;  
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;  
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;  
KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.  
XX

OS Homo sapiens.  
XX

PN US6251864-B1.  
XX

PD 26-JUN-2001.  
XX

PF 01-MAR-2000; 2000US-00516704.  
XX

PR 07-JUN-1995; 95US-00478128.  
XX

PR 07-JUN-1995; 95US-00485301.  
XX

PR 07-JUN-1996; 96WO-US009623.  
XX

PR 15-AUG-1996; 96US-00699027.  
XX

PA (GLAXO ) GLAXO GROUP LTD.  
XX

PI Dower WJ, Barrett RM, Cwiria SE, Gates CM, Schatz PJ;  
XX Balasubramanian P, Wagstrom CR, Hendren RW, Depirnce RB, Podduturi S;  
XX yin Q;  
XX

DR WPI; 2001-564142/63.  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

PT  
XX

CC be used for in vitro expansion of megakaryocytes and their committed  
XX progenitors alone or in conjunction with additional cytokines  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAAARA 14  
| | | | | | | | | |  
| | | | | | | | | |  
Db 2 IEGPTLRQWLAAARA 15

Search completed: September 1, 2005, 16:12:15  
Job time : 65.3597 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: September 1, 2005, 15:57:33 ; Search time 10.6763 Seconds  
(without alignments)  
126.171 Million cell updates/sec

Title: US-10-083-768-13

Perfect score: 73

Sequence: 1 IEGPTRQWLARA 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 100 summaries

Database :

PIR 79: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.5	67.8	333	2 A36925	transcription acti
2	47	64.4	296	2 AG0147	probable membrane
3	46	63.0	306	2 D70601	UTP-glucose-1-phos
4	44	60.3	200	2 T23485	hypothetical prote
5	44	60.3	207	2 T37464	probable glutathio
6	44	60.3	536	1 SYECB	2,3-dihydroxybenzo
7	44	60.3	536	2 E85558	2,3-dihydroxybenzo
8	44	60.3	536	2 A99708	probable dimethyla
9	43	58.9	285	2 G71337	conserved hypochet
10	43	58.9	683	2 B71325	hypothetical prote
11	42	57.5	473	2 E84853	pol polyprotein -
12	42	57.5	1019	2 T11560	probable phosphopa
13	41	56.2	195	2 F91171	hypothetical 21.8k
14	41	56.2	195	2 F86017	ABC transporter, A
15	41	56.2	195	2 S47694	UTP-glucose-1-phos
16	41	56.2	249	2 E87575	glyceraldhyde-3-p
17	41	56.2	306	2 T45453	glyceraldhyde-3-p
18	41	56.2	326	2 C24430	glyceraldhyde-3-p
19	41	56.2	336	1 DEPPG	glyceraldhyde-3-p
20	41	56.2	337	2 A35080	glyceraldhyde-3-p
21	41	56.2	338	1 DEIS3C	glyceraldhyde-3-p
22	41	56.2	338	2 J01287	conserved hypochet
23	41	56.2	719	2 B95325	topoisomerase IV c
24	41	56.2	750	2 A97501	topoisomerase IV c
25	41	56.2	750	2 AE2719	SRM protein - Str
26	40	54.8	239	2 S25204	hypothetical prote
27	40	54.8	463	2 S27491	probable permease
28	40	54.8	530	2 A81958	ABC transporter, p
29	40	54.8	531	2 E81015	

30	40	54.8	656	2	S30484	pol polyprotein -
31	40	54.8	656	2	S30483	pol polyprotein -
32	40	54.8	721	2	A39707	erythrocyte membra
33	40	54.8	1123	2	T51517	telomerase reverse
34	40	54.8	1712	1	CGHJ28	collagen alpha 2(I
35	39.5	54.1	325	2	A84326	hypothetical prote
36	39.5	54.1	131	2	S74539	hypothetical prote
37	39	53.4	217	2	S46354	pol polyprotein -
38	39	53.4	267	2	I40327	baf protein - Bord
39	39	53.4	331	2	B48445	glyceraldehyde-3-p
40	39	53.4	331	2	A72514	hypothetical prote
41	39	53.4	400	2	C87021	serine-threonine p
42	39	53.4	600	2	C83221	transport protein
43	39	53.4	791	2	A82291	c-di-GMP phosphodi
44	39	53.4	1034	1	GNLJCA	HIV-1 retropepsin
45	39	53.4	1035	1	GNLJGG	HIV-1 retropepsin
46	39	53.4	1036	1	GNLJG2	HIV-1 retropepsin
47	39	53.4	1055	1	GNLJST	pol polyprotein -
48	39	53.4	1055	2	S53092	pol polyprotein -
49	39	53.4	1058	2	S08436	hypothetical prote
50	39	53.4	1058	2	T13423	hypothetical prote
51	38	52.1	134	2	B73468	transcription regu
52	38	52.1	197	2	G82973	probable oxidoredu
53	38	52.1	246	2	AH0190	glyceraldehyde-3-p
54	38	52.1	247	2	PQ0178	glyceraldehyde-3-p
55	38	52.1	295	2	T07730	glyceraldehyde-3-p
56	38	52.1	297	2	B87109	integrase/recombin
57	38	52.1	311	1	RGECK	regulatory protein
58	38	52.1	311	1	AH0867	transcription acti
59	38	52.1	311	2	C85936	positive regulator
60	38	52.1	311	2	H91090	hypothetical prote
61	38	52.1	314	2	H70723	hypothetical prote
62	38	52.1	337	1	DEBPG	glyceraldehyde-3-p
63	38	52.1	337	1	DEBSKG	glyceraldehyde-3-p
64	38	52.1	337	1	DEBSGM	glyceraldehyde-3-p
65	38	52.1	337	1	DEZMGC	hypothetical prote
66	38	52.1	339	2	A83358	A/G-specific adeni
67	38	52.1	339	2	B38535	adenine glycosylas
68	38	52.1	350	2	H85953	adenine glycosylas
69	38	52.1	350	2	E91108	hypothetical prote
70	38	52.1	360	2	S38570	hypothetical prote
71	38	52.1	469	2	AD1926	sensor histidine k
72	38	52.1	589	2	F87626	hypothetical prote
73	38	52.1	635	2	A87433	hypothetical prote
74	38	52.1	816	2	A71006	hypothetical prote
75	38	52.1	904	2	C70559	probable polA prot
76	38	52.1	1155	2	AC2426	adenylate cyclase
77	37.5	51.4	436	2	JC4742	transposase - Cory
78	37	50.7	151	2	S63748	HIV-1 retropepsin
79	37	50.7	165	2	F87542	hypothetical prote
80	37	50.7	234	2	PQ0179	glyceraldehyde-3-p
81	37	50.7	305	2	A24159	glyceraldehyde-3-p
82	37	50.7	335	2	S29813	glyceraldehyde-3-p
83	37	50.7	337	1	DEBHG	glyceraldehyde-3-p
84	37	50.7	337	2	S42479	glyceraldehyde-3-p
85	37	50.7	337	2	T02723	glyceraldehyde-3-p
86	37	50.7	337	2	T02722	glyceraldehyde-3-p
87	37	50.7	338	1	DENDG	glyceraldehyde-3-p
88	37	50.7	338	2	T06781	glyceraldehyde-3-p
89	37	50.7	341	1	DEJMG	glyceraldehyde-3-p
90	37	50.7	341	2	T08147	glyceraldehyde-3-p
91	37	50.7	341	2	AG0195	probable exported
92	37	50.7	352	2	G83636	conserved hypochet
93	37	50.7	391	2	T36739	hypothetical prote
94	37	50.7	407	2	A86298	hypothetical prote
95	37	50.7	422	2	F96826	hypothetical prote
96	37	50.7	433	2	S51837	glyceraldehyde-3-p
97	37	50.7	433	2	S51836	glyceraldehyde-3-p
98	37	50.7	438	2	G87337	membrane protein,
99	37	50.7	480	2	H84747	probable steroid d
100	37	50.7	486	2	B86411	protein F3M18.4 [1



A:Residues: 1-207 <TRAM>  
 A:Cross-references: UNIPROT:O16116; EMBL:AF010241; PDB:AA05419.1  
 A:Experimental source: strain Bristol N2  
 C:Genetic:  
 A:Gene: GST3  
 C:Superfamily: glutathione transferase  
 C:Keywords: transferase

Query Match 60.3%; Score 44; DB 2; Length 207;  
 Best Local Similarity 61.5%; Pred. No. 5.1;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 13  
 DB 190 IETPKLEWLAKR 202

RESULT 6  
 SYCEB  
 2,3-dihydroxybenzoate-[carrier protein] ligase (EC 6.2.1.-) ente - Escherichia coli (str N1/Alternae names: 2,3-dihydroxybenzoate-AMP ligase [misnomer]; dihydroxybenzoic acid-ad C/Species: Escherichia coli  
 C/Date: 31-Dec-1989 #sequence\_revision 21-Nov-1997 #text\_change 09-Jul-2004  
 C/Accession: H64792; A48308; A32047; I41058; S08076  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: H64792  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-536 <BLAT>  
 A:Cross-references: UNIPROT:P10378; GB:AE000165; GB:T00096; NID:G1786808; PDB:AACT3695.  
 A:Experimental source: strain K-12, substrain MG1655  
 R:Staab, J.F.; Elkins, M.F.; Barthart, C.F.  
 FEMS Microbiol. Lett. 59, 15-19, 1989  
 A:Title: Nucleotide sequence of the Escherichia coli ente gene.  
 A:Reference number: A48308; MUID:89230355; PMID:2525505  
 A:Note: In Medline 89290355 this citation is erroneously given as volume 50 rather than 59.  
 A:Accession: A48308  
 A:Molecule type: DNA  
 A:Residues: 1366, 'GCRKSTAA', 379-536 <STA>  
 A:Cross-references: GB:M27490; EMBL:X15058; NID:G41345; PDB:CAA33158.1; PID:G41346  
 R:Li, J.; Duncan, K.; Walsh, C.T.  
 J. Bacteriol. 171, 791-798, 1989  
 A:Title: Nucleotide sequence of a cluster of Escherichia coli enterobactin biosynthesis A:Reference number: A91904; MUID:89123155; PMID:2521622  
 A:Accession: A32047  
 A:Molecule type: DNA  
 A:Residues: 393-536 <LIU>  
 A:Cross-references: GB:M24148; NID:G304949; PDB:AAA16101.1; PID:G450380  
 C/Comment: The enzymatic steps in the condensation of L-serine and 2,3-dihydroxybenzoic ty is based on the recognized homology with 4-coumarate-CoA ligase and by analogy with C/Comment: The formation of 2,3-dihydroxybenzoyl-AMP has been observed. The rapid reaction carrier protein) to release AMP, has also been observed.  
 C:Genetic:  
 A:Gene: ente  
 A:Map position: 14 min  
 C:Function:  
 A:Description: catalyzes the formation of 2,3-dihydroxybenzoyl-[carrier protein], AMP and A:Pathway: enterobactin biosynthesis  
 A:Note: this is one component of a membrane-bound multienzyme complex that catalyzes the for transport into the cell  
 C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology  
 C:Keywords: acid-thiol ligase; enterobactin biosynthesis; membrane-associated complex  
 P:69-526/Domain: acetate-CoA ligase homology <ACL>

Query Match 60.3%; Score 44; DB 1; Length 536;  
 Best Local Similarity 57.1%; Pred. No. 13;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IEGPTLRQWLAAAR 14

DB 521 VDKKQRLQWLASRA 534

RESULT 7  
 E85558  
 2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substra C/Species: Escherichia coli  
 C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 01-Mar-2002  
 C/Accession: E85558  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimianta, E.; Potamousis, K.; Apodaca Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: E85558  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-536 <STO>  
 A:Cross-references: GB:AE005174; NID:G12513487; PDB:AA054929.1; GSPDB:GN00145; UMGDB:20 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetic:  
 A:Gene: ente  
 C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;  
 Best Local Similarity 57.1%; Pred. No. 13;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 14  
 DB 521 VDKKQRLQWLASRA 534

RESULT 8  
 A99708  
 2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substra C/Species: Escherichia coli  
 C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
 C/Accession: A99708  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: A99708  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-536 <HAY>  
 A:Cross-references: UNIPROT:Q8XEVJ; GB:BA000007; PDB:BAR34056.1; PID:G13360091; GSPDB: A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetic:  
 A:Gene: ECG0633  
 C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;  
 Best Local Similarity 57.1%; Pred. No. 13;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 14  
 DB 521 VDKKQRLQWLASRA 534

RESULT 9  
 G71337  
 probable dimethyladenosine transferase (KsgA) - syphilis spirochete C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C/Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
 C/Accession: G71337  
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi rson, J.; Khatalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998  
 A>Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
 A|Reference number: A71250; M01D:98332770; PMID:9665876  
 A|Accession: G71337  
 A|Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A|Molecule type: DNA  
 A|Residues: 1-285 <COL>  
 A|Cross-references: UNIPROT:O83357; GB:AE001213; GB:AE000520; NID:G3322606; PIDN:AAC6532  
 A|Experimental source: strain Nichols  
 C|Genetics:  
 A|Gene: TP0337  
 C|Superfamily: rRNA (adenine-N6--methyl)transferase  
  
 Query Match                      58.9%;      Score 43;      DB 2;      Length 285;  
 Best Local Similarity      64.3%;      Pred. No. 10;  
 Matches      9;      Conservative      1;      Mismatches      4;      Indels      0;      Gaps      0;  
  
 Oy                      1      IEGPTLRQWLAARA      14  
                          |||||      :|||      |||  
 Db                      98      IEGDVLQGMHAAA      111

RESULT 10  
B71325  
conserved hypothetical protein TP0421 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: B71325  
R:Reisser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwiniłł, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ueterbach, T.; McDaniel, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
S:Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A12520, MUID:98332770, PMID:9665876  
A:Accession: B71325  
A:Status: preliminary, nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-683 <COL>  
A:Cross-references: UNIPROT:O83436; GB:AE001220; GB:AE000520, NID:G3322705; PIDDN:PA6C540  
A:Experimental source: strain Nichols  
C:Genetics:  
C:Gene: TP0421

```

Query Match      58.9%  Score 43;  DB 2;  Length 683;
Best Local Similarity 69.2%  Pred. No. 25;
Matches 9;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;

OY      1  IEGPTRLQWLAAAR 13
      |||  |||  |||
      |||  |||  |||
Db      89  IEGAAALHQMGAAR 101

RESULT 11
E84853
hypothetical protein At2g42400 [imported] - Arabidopsis thaliana
CISpecies: Arabidopsis thaliana (mouse-ear cress)
CDate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
CAccession: E84853
RLin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Native 402, 761-766, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A|Reference number: A64420; MUID:20083487; PMID:10617197
A|Accession: E84853
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-473 <STO>
A|Cross-references: UNIPROT:O9SLB9; GB:AE002093; NID:g4567312; PIDN:AA023723.1; GSPDB:GN
C|Genetics:
A|Gene: At2g42400
A|Map position: 2

```

Query Match	57.5%	Score 42;	DB 2;	Length 473;
Best Local Similarity	60.0%	Pred. No. 26;		
Matches	6;	Conservative	3;	Mismatches
			1;	Indels
				Gaps
Oy	1	IEGPTLRQWL	10	
	:			
Db	343	VEGETIREWL	352	

RESULT 12  
T11560  
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)  
!Species: simian immunodeficiency virus SIVsm

A:arvady: 5:chain\_b3  
 C:date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:accession: J11560  
 R:Ritsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.;  
 J. Virol. 71, 1608-1620, 1997  
 A:title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficient  
 Reference number: Z17285; MUID:97151152; PMID:8995668

A:Accession: [U02200](#)  
A:Status: preliminary; translated from GB/EMBL/DBD1  
A:Molecule type: DNA  
A:Residues: 1-1019 <HR>  
A:Cross-references: UNIPROT:P89154; EMBL:U72748; NID:G1695908; PIDN:AACS6559.1; PID:G165  
A:Genetics:  
A:Gene: pol  
C:Superfamily: pol polypeptide  
C:Keywords: AIDS; immunodeficiency

Query Match	57.5%	Score 42;	DB 2;	Length 1019;
Best Similarity	87.5%	Fred. No. 56;		
Best Local	7;	Conservative	0;	Mismatches
Matches			1;	Indels
			0;	Gaps
0Y	2	EGPTLR0W	9	
DB	184	EGPTLR0W	191	

RESULT 13  
P91171  
probable phosphopantetheinyltransferase [imported] - Escherichia coli (strain O157:H7, {  
C)Species: Escherichia coli  
C)Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C)Accession: F91171  
R)Hayashi, T.; Makino, K.; Ohmishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.  
Sasagawa, N.; Yasunaga, T.; Kuhrara, S.; Shiba, T.; Hartfort, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A)Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A)Reference number: A9629; MUID:21156231; PMID:11258796  
A)Accession: F91171  
A)Molecule type: DNA  
A)Status: preliminary  
A)Residues: 1-795 <HAV>  
A)Cross-references: UNIPROT:Q8Y5U4; GB:BA000007; PIDD:BA837765.1; PID:g13363816; GSPDB:(  
A)Experimental source: strain O157:H7, Substrain RMD 0509952  
C)Genetics:  
A)Gene: ECe4342

Query Match	Score 41;	DB 2;	Length 195;
Best Similarity	56.2%		
Best Local	53.8%	Pred. No. 15;	
Matches	7;	Conservative	2; Mismatches 4; Indels 0; Gaps 0
QY	2	EGPTLRQWTLARA	14
	:	:	:
	:	:	:
DB	27	OGPRRRRLAGRA	39

```

RESULT 14
F86017
Probable phosphopantetheinyltransferase [Imported] - Escherichia coli (strain O157:H7,
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence revision 16-Feb-2001 #text change 09-Jul-2004

```

C/Accession: F86017  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; DiMantana, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A/Accession: F86017  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-195 <STO>  
 A:Cross-references: UNIPROT:Q8X5U4; GB:AE005174; NID:g12518155; PIDN:ANG58602.1; GSPDB:C  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C/Genetics:  
 A:Gene: 24867

Query Match 56.2%; Score 41; DB 2; Length 195;  
 Best Local Similarity 53.8%; Pred. No. 15;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 EGPTLRQWLARA 14  
 :|||:|||||  
 Db 27 QGPRRRRWLAGRA 39

RESULT 15  
 S47694  
 hypothetical 21.8K protein (feyr-nika intergenic region) - *Escherichia coli* (strain K-12  
 N:Alternate names: hypothetical protein o195  
 C/Species: *Escherichia coli*  
 C/Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 09-Jul-2004  
 C/Accession: S47694; F85144  
 R:Plunkett, G.  
 submitted to the EMBL Data Library, March 1994  
 A/Reference number: S47666  
 A/Accession: S47694  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-195 <PLU>  
 A:Cross-references: UNIPROT:P37623; EMBL:U00039; NID:g466582; PIDN:AA18450.1; PID:g4666  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
 A.; Rose, D.J.; Mau, B.; Snao, Y.  
 Science 277, 1453-1462, 1997  
 A>Title: The complete genome sequence of *Escherichia coli* K-12.  
 A:Reference number: A64720; MUID:9742617; PMID:9278503  
 A/Accession: F85144  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-195 <BLAT>  
 A:Cross-references: GB:AE000423; GB:U00096; NID:g1789880; PIDN:AACT6500.1; PID:g1789886;  
 A:Experimental source: strain K-12, substrain MG1655  
 C/Genetics:  
 A:Gene: ynhu

Query Match 56.2%; Score 41; DB 2; Length 195;  
 Best Local Similarity 53.8%; Pred. No. 15;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 EGPTLRQWLARA 14  
 :|||:|||||  
 Db 27 QGPRRRRWLAGRA 39

RESULT 16  
 E87575  
 ABC transporter, ATP-binding protein CC2634 [imported] - *Caulobacter crescentus*  
 C/Species: *Caulobacter crescentus*  
 C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C/Accession: E87575  
 R:Neuman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eissen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A>Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A/Accession: E87575  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-249 <STO>  
 A:Cross-references: UNIPROT:Q9A535; GB:AE005673; NID:g13424211; PIDN:AAK24601.1; GSPDB:  
 C/Genetics:  
 A:Gene: CC2634

Query Match 56.2%; Score 41; DB 2; Length 249;  
 Best Local Similarity 58.3%; Pred. No. 20;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EGPTLRQWLAR 13  
 :|||:|||||  
 Db 76 QAPTLAPWLSAR 87

RESULT 17  
 T45453  
 UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - *Mycobacter*  
 C/Species: *Mycobacterium leprae*  
 C/Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004  
 C/Accession: T45453  
 R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, February 1998  
 A/Reference number: Z22967  
 A/Accession: T45453  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-306 <JMB>  
 A:Cross-references: UNIPROT:Q9Z5G1; EMBL:AL035500; PIDN:CAB36696.1  
 A:Experimental source: cosmid U373  
 C/Genetics:  
 A:Note: galU  
 C:Superfamily: *Escherichia coli* UTP-glucose-1-phosphate uridylyltransferase  
 C:Keywords: nucleotidyltransferase

Query Match 56.2%; Score 41; DB 2; Length 306;  
 Best Local Similarity 63.6%; Pred. No. 24;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 GPTLRQWLAR 13  
 :|||:|||||  
 Db 290 GPDLRKWLVER 300

RESULT 18  
 C24430  
 glyceraldehyde-3-phosphate dehydrogenase (NADP) (phosphorylating) (EC 1.2.1.13) C. cyto  
 C/Species: *Nicotiana tabacum* (common tobacco)  
 C/Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 09-Jul-2004  
 C/Accession: C24430  
 R:Shih, M.C.; Lazar, G.; Goodman, H.M.  
 Cell 47, 73-80, 1986  
 A>Title: Evidence in favor of the symbiotic origin of chloroplasts: primary structure a  
 A:Reference number: A90888; MUID:87002494; PMID:3757034  
 A/Accession: C24430  
 A:Molecule type: mRNA  
 A:Residues: 1-326 <SHI>  
 A:Cross-references: UNIPROT:P09094; GB:M14419; NID:g170240; PIDN:AAA34077.1; PID:g170240  
 C/Genetics:  
 A:Gene: gapC  
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase  
 C:Keywords: cytosol; NADP; oxidative phosphorylation; oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 326;  
 Best Local Similarity 35.7%; Pred. No. 26;  
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 IEPTLRQWLARA 14  
 :|||:|||||  
 Db 179 VDGSMDWAGRA 192



RESULT 19  
DEPZG  
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - parsley  
C/Species: Petroselinum crispum (parsley)  
C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C/Accession: S18484  
R/Martin, W.; Gierl, A.; Saedler, H.  
Nature 339, 46-48, 1989  
A/Title: Molecular evidence for pre-Cretaceous angiosperm origins.  
A/Reference number: S17991  
A/Accession: S18484  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-336 <MAR>  
A/Cross-references: UNIPROT:P26519; EMBL:X60344; NID:G20548; PIDN:CAA42902.1; PID:G20549  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1991  
C/Superfamily: glyceralddehyde-3-phosphate dehydrogenase  
C/Keywords: gluconeogenesis; glycolysis; homocitrate; NAD; oxidoreductase  
F/4-34/Region: beta-alpha-beta NAD nucleotide-binding fold  
F/153,180/Active site: Cys, His #status predicted

Query Match 56.2%; Score 41; DB 1; Length 336;  
Best Local Similarity 35.7%; Pred. No. 27;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
::|::|::|::|  
DB 189 VDGPSMKDMRGRA 202

RESULT 20  
A35080  
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - common ice plant  
C/Species: Mesembryanthemum crystallinum (common ice plant)  
C/Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 09-Jul-2004  
C/Accession: A35080  
R/Ostrem, J.A.; Vernon, D.M.; Bohnert, H.J.  
J. Biol. Chem. 265, 3497-3502, 1990  
A/Title: Increased expression of a gene coding for NAD:glyceralddehyde-3-phosphate dehydrogenase.  
A/Reference number: A35080; MUID:90154012; PMID:2203458  
A/Accession: A35080  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-337 <OST>  
A/Cross-references: UNIPROT:P17878; GB:J05223; NID:G167263; PIDN:AAA3033.1; PID:G167264  
C/Superfamily: glyceralddehyde-3-phosphate dehydrogenase  
C/Keywords: oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 337;  
Best Local Similarity 35.7%; Pred. No. 27;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
::|::|::|::|  
DB 190 VDGPSMKDMRGRA 203

RESULT 21  
DEIS3C  
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - white mustard  
N/Alternate names: triosephosphate dehydrogenase  
C/Species: Sinapis alba (white mustard)  
C/Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C/Accession: A24796  
R/Martin, W.; Cerff, R.  
Eur. J. Biochem. 199, 323-331, 1986  
A/Title: Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast  
A/Reference number: A24796; MUID:87004643; PMID:3530755  
A/Accession: A24796  
A/Molecule type: mRNA

A/Residues: 1-338 <MAR>  
A/Cross-references: UNIPROT:P04796; GB:X04301; NID:G21142; PIDN:CAA27844.1; PID:G21143  
C/Superfamily: glyceralddehyde-3-phosphate dehydrogenase  
C/Keywords: gluconeogenesis; glycolysis; homocitrate; NAD; oxidoreductase  
F/2-338/Product: glyceralddehyde-3-phosphate dehydrogenase #status experimental <MAR>  
F/7-37/Region: beta-alpha-beta NAD nucleotide-binding fold  
F/156,183/Active site: Cys, His #status predicted

Query Match 56.2%; Score 41; DB 1; Length 338;  
Best Local Similarity 35.7%; Pred. No. 27;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
::|::|::|::|  
DB 192 VDGPSMKDMRGRA 205

RESULT 22  
JQ1287  
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - Arabidopsis thaliana (mouse-ear cress)  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C/Accession: JQ1287; J50614  
R/Shih, M.C.; Heinrich, P.; Goodman, H.M.  
Gene 104, 133-138, 1991  
A/Title: Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic  
A/Reference number: JQ1285; MUID:92009205; PMID:1916285  
A/Accession: JQ1287  
A/Molecule type: DNA  
A/Residues: 1-338 <SHI>  
A/Cross-references: UNIPROT:P25858; GB:M64119; NID:G166709; PIDN:AAA32796.1; PID:G166710  
A/Accession: J50614  
A/Molecule type: mRNA  
A/Residues: 1-338 <SHI>  
A/Cross-references: GB:M64116; NID:G166705; PIDN:AAA32794.1; PID:G166706  
A/Experimental source: leaf  
C/Genetic: 8  
A/Genetic: gacc  
A/Map position: 3 0.0cm  
A/Intons: 2/1; 12/1; 45/3; 84/2; 117/3; 167/2; 187/1; 267/2  
C/Superfamily: glyceralddehyde-3-phosphate dehydrogenase  
C/Keywords: cytosol; oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 338;  
Best Local Similarity 35.7%; Pred. No. 27;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14  
::|::|::|::|  
DB 192 VDGPSMKDMRGRA 205

RESULT 23  
B95325  
conserved hypothetical protein Sma0937 [imported] - Sinorhizobium meliloti (strain 1021)  
C/Species: Sinorhizobium meliloti  
C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C/Accession: B95325  
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow, J.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melillo  
A/Reference number: A95262; MUID:21596509; PMID:11481432  
A/Accession: B95325  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-719 <KUR>  
A/Cross-references: UNIPROT:Q92ZH9; GB:AE006469; PIDN:AAK65164.1; PID:G14523607; GSPDB: A/Experimental source: strain 1021, megaplasmid pSymA  
R/Gallbert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001



A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, M.; hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMO937  
A:Genome: plasmid

Query Match 56.2%; Score 41; DB 2; Length 719;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEPTLQWMLAAR 13  
Db 71 LDDEVRQWMLTAK 83

RESULT 24  
A97501  
topoisomerase iv chain a [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: A97501  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97559; MUID:21608551; PMID:11743194  
A:Accession: A97501  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-750 <KUR>  
A:Cross-references: UNIPROT:Q8UG82; GB:AB007869; PIDN:AAK86962.1; PID:g15156198; GSPDB:Q

C:Genetics:  
A:Gene: AGR\_C\_2144  
A:Map position: circular chromosome  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase  
Query Match 56.2%; Score 41; DB 2; Length 750;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPTLRQWMLAAR 14  
Db 721 GBLRQWMLAAR 732

RESULT 25  
AB2719

topoisomerase IV subunit A parC [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AB2719

R:Wood, D.W.; Setudal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavln, T.; Levy, R.; Li, M.; Mclellan, S.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; B. W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AB2719  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-750 <KUR>

A:Cross-references: UNIPROT:Q8UG82; GB:AB008688; PIDN:AA42171.1; PID:g17739560; GSPDB:Q  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: parC  
A:Map position: circular chromosome  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match 56.2%; Score 41; DB 2; Length 750;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPTLRQWMLAAR 14  
Db 721 GBLRQWMLAAR 732

RESULT 26

S25204  
srnx protein - Streptomyces ambofaciens

C:Species: Streptomyces ambofaciens  
C:Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 09-Jul-2004  
C:Accession: S25204; S21599  
R:Geistlich, M.; Losick, R.; Turner, J.R.; Rao, R.N.  
Mol. Microbiol. 6, 2019-2029, 1992  
A:Title: Characterization of a novel regulatory gene governing the expression of a pol  
A:Reference number: S25202; MUID:92374852; PMID:1508047  
A:Accession: S25204  
A:Molecule type: DNA  
A:Residues: 1-239 <GBI>  
A:Cross-references: UNIPROT:Q00510; EMBL:X63451; NID:946699; PIDN:CAA45052.1; PID:g46570

C:Genetics:  
A:Gene: srnx  
F:39-139/Domain: bioc homology <BIOC>  
Query Match 54.8%; Score 40; DB 2; Length 239;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEPTLQWMLAAR 14  
Db 63 VSGLESEWMAAR 76

RESULT 27

S27491  
hypothetical protein A - Bacillus firmus

C:Species: Bacillus firmus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S27491  
R:Quirk, P.G.; Krulwich, T.A.  
submitted to the EMBL Data Library, October 1991  
A:Reference number: S27490  
A:Accession: S27491  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-463 <QHI>  
A:Cross-references: UNIPROT:P30267; GB:L02548; EMBL:M74194; NID:g143118; PIDN:AAA22559.

Query Match 54.8%; Score 40; DB 2; Length 463;  
Best Local Similarity 61.5%; Pred. No. 54;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EGPTLRQWMLAAR 14  
Db 296 EGKTSRQWMLAAR 308

RESULT 28

AB1958  
probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)

C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C:Accession: AB1958  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More; Hojrova, S.; Jagers, K.; Leather, S.; Moutle, S.; Mungall, K.; Quail, M.A.; Rajandream; Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: AB1775; MUID:20222556; PMID:10761919  
A:Accession: AB1958  
A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-530 <PAR>  
 A:Cross-references: UNIPROT:Q9JWB3; GB:AL162753; GB:AL157959; NID:G7379120; PIDN:CAH8371  
 A:Experimental source: serogroup A, strain 22491  
 C:Genetics:  
 A:Gene: NMB0414

Query Match 54.8%; Score 40; DB 2; Length 530;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGETLRQWLA 11  
 ||| ||| |||  
 Db 190 IEMPVLRPWLA 200

## RESULT 29

AB0105  
 ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C:Accession: E81015  
 R:Retelid, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, V.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: AB0100; MUID:2015755; PMID:10710307  
 A:Accession: E81015  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-531 <RET>  
 A:Cross-references: UNIPROT:Q9JX19; GB:AE002552; GB:AE002098; NID:G7227279; PIDN:AAE4234  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB2026

Query Match 54.8%; Score 40; DB 2; Length 531;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGETLRQWLA 11  
 ||| ||| |||  
 Db 191 IEMPVLRPWLA 201

## RESULT 30

S30484  
 pol polyprotein - human immunodeficiency virus type 2  
 C:Species: human immunodeficiency virus type 2, HIV-2  
 C:Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001  
 C:Accession: S30484  
 R:Go, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
 A:Reference number: S30460  
 A:Accession: S30484  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-656 <GAO>  
 A:Cross-references: EMBL:M87114  
 C:Superfamily: pol polyprotein

Query Match 54.8%; Score 40; DB 2; Length 656;  
 Best Local Similarity 66.7%; Pred. No. 77;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGETLRQW 9  
 :||| |||  
 Db 29 MDGPKLRQW 37

## RESULT 31

S30483  
 pol polyprotein - human immunodeficiency virus type 2  
 C:Species: human immunodeficiency virus type 2, HIV-2  
 C:Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 23-Mar-2001  
 C:Accession: S30483  
 R:Go, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa  
 A:Reference number: S30460  
 A:Accession: S30483  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-656 <GAO>  
 A:Cross-references: EMBL:M87111  
 C:Superfamily: pol polyprotein

Query Match 54.8%; Score 40; DB 2; Length 656;  
 Best Local Similarity 66.7%; Pred. No. 77;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGETLRQW 9  
 :||| |||  
 Db 29 MDGPKLRQW 37

## RESULT 32

A39707  
 erythrocyte membrane band 4.2 protein - human  
 N:Alternate names: pallidin  
 N:Contains: erythrocyte membrane band 4.2 protein, long splice form; erythrocyte membra  
 C:Species: Homo sapiens (man)  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: A39707; A34865; B34865; A34883  
 R:Korsgren, C.; Cohen, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 4840-4844, 1991  
 A:Title: Organization of the gene for human erythrocyte membrane protein 4.2: structure  
 A:Reference number: A39707; MUID:91271288; PMID:2052563  
 A:Accession: A39707  
 A:Molecule type: DNA  
 A:Residues: 1-721 <KOR1>  
 A:Cross-references: UNIPROT:P16452; GB:L06519; NID:G306738; PIDN:AAA52385.1; PID:G30674  
 A:Experimental source: cell type erythrocyte; tissue type peripheral blood; tissue lib  
 R:Sun, L.A.; Chien, S.; Chang, L.S.; Lambert, K.; Blise, S.A.; Bouhasstra, E.E.; Nagel  
 Proc. Natl. Acad. Sci. U.S.A. 87, 955-959, 1990  
 A:Title: Molecular cloning of human protein 4.2: a major component of the erythrocyte m  
 A:Reference number: A34865; MUID:90138995; PMID:1689063  
 A:Accession: A34865  
 A:Molecule type: mRNA

A:Residues: 1364, 'KRGLPC', 371-379, 'H', 381-405, 'L', 407-721 <SUN1>  
 A:Cross-references: GB:M30647; NID:G189433; PIDN:AAA56401.1; PID:G189434  
 A:Accession: B34865  
 A:Molecule type: mRNA  
 A:Residues: 1-3, 34-721 <KOR2>  
 A:Cross-references: GB:M30646; NID:G189435; PIDN:AAA56402.1; PID:G189436  
 A:Experimental source: isolate Sickie cell patient; cell type reticulocyte  
 A>Note: parts of this sequence were determined by protein sequencing  
 R:Korsgren, C.; Lawler, J.; Lambert, S.; Specht, D.; Cohen, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 613-617, 1990  
 A:Title: Complete amino acid sequence and homologies of human erythrocyte membrane prot  
 A:Reference number: A34883; MUID:90138879; PMID:2300550  
 A:Accession: A34883  
 A:Molecule type: mRNA  
 A:Residues: 1-3, 34-721 <KOR2>  
 A:Cross-references: GB:M29399; NID:G182083; PIDN:AAA5798.1; PID:G182084  
 C:Comment: This protein is a major constituent of the erythrocyte membrane. It apparent  
 C:Genetics:  
 A:Gene: GDB:EBB42; PA  
 A:Cross-references: GDB:127385; OMIM:177070  
 A:Map position: 15q15-15q15  
 C:Superfamily: protein-glutamine gamma-glutamyltransferase  
 C:Keywords: alternative splicing; blocked amino end; glycoprotein; lipoprotein; myristic  
 F:2-721/Product: erythrocyte membrane band 4.2 protein, long splice form #status predic

F,2,3,34-721/Product: erythrocyte membrane band 4.2 protein, short splice form #status F  
 F,298-316/Domain: transmembrane #status predicted <TRM>  
 F,518-520/Region: cell attachment (R-G-D) motif  
 F,2/Modified site: myristylated amino end (GIY) (in mature form) #status predicted  
 F,103,420,447,529,604,705/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F,218/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted

Query Match 54.8%; Score 40; DB 2; Length 721;  
 Best Local Similarity 70.0%; Pred. No. 85;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 PTLRQMLAA 13  
 |||||  
 Db 280 PTLRQMLTGR 289

RESULT 33  
 T51517  
 telomerase reverse transcriptase - Arabidopsis thaliana  
 N/Alternate names: protein F5E19\_190  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
 C/Accession: T51517  
 R/Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tebata, S.; New  
 submitted to the Protein Sequence Database, August 2000  
 A/Reference number: 225394  
 A/Accession: T51517  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1123 <SAT>  
 A/Cross-references: UNIPROT:Q9SPU7; EMBL:AL391147  
 C/Experimental source: cultivar Columbia; BAC clone F5E19  
 C/Genetics:  
 A/Map position: '5  
 A/Introns: 100/3; 125/3; 147/3; 165/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2  
 A/Note: F5E19\_190

Query Match 54.8%; Score 40; DB 2; Length 1123;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 IEPTLRQMLAA 12  
 ::|||:::  
 Db 200 VQPTKQWLSS 211

RESULT 34  
 CGH028  
 collagen alpha 2(IV) chain precursor - human  
 N/Alternate names: procollagen alpha 2(IV) chain  
 C/Species: Homo sapiens (man)  
 C/Date: 07-Jun-1990 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
 C/Accession: A32024; S00007; S02624; S00246; S17678; S16911; B32117; S16877; S00165; S39  
 R/Hoslika, S.L.; Trygsvaen, K.  
 J. Biol. Chem. 263, 19488-19493, 1988  
 A/Title: The complete primary structure of the alpha2 chain of human type IV collagen an  
 A/Reference number: A32024; MUID:8906769; PMID:3198637  
 A/Accession: A32024  
 A/Molecule type: mRNA  
 A/Residues: 1-1712 <HOS1>  
 A/Cross-references: UNIPROT:P08572; EMBL:J04210; EMBL:X05610; GB:M20753; NID:G29550; PID  
 R/Hoslika, S.L.; Kurkinen, M.; Trygsvaen, K.  
 FEBS Lett. 216, 281-286, 1987  
 A/Title: Nucleotide sequence coding for the human type IV collagen alpha-2 chain cDNA re  
 ated region.  
 A/Reference number: S00007; MUID:87219158; PMID:3582677  
 A/Accession: S00007  
 A/Molecule type: mRNA  
 A/Residues: 1254-1398, 'V', 1400-1712 <HOS2>  
 A/Cross-references: EMBL:J04210; EMBL:X05610; GB:M20753; NID:G29550; PIDN:CAA29076.1; PI  
 R/Hoslika, S.L.; Trygsvaen, K.  
 FEBS Lett. 224, 297-305, 1987

A/Title: Extensive structural differences between genes for the alpha(1) and alpha(2) c  
 A/Reference number: S02624; MUID:88083553; PMID:2826228  
 A/Accession: S02624  
 A/Status: not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 1347-1350;1377-1383;1426-1432;1465-1471;1529-1535;1625-1630 <HOS3>  
 A/Note: complete nucleotide sequence not shown  
 R/Brazel, D.; Pollner, R.; Oberhauser, I.; Kuehn, K.  
 Eur. J. Biochem. 172, 35-42, 1988  
 A/Title: Human basement membrane collagen (type IV): the amino acid sequence of the al  
 A/Reference number: S00246; MUID:88151998; PMID:3345760  
 A/Accession: S00246  
 A/Molecule type: mRNA  
 A/Residues: 1-682, 'G', 684-1043 <BR>  
 A/Cross-references: EMBL:X05562; NID:G30075; PIDN:CAA29076.1; PID:G30076  
 R/Oberhauser, I.  
 submitted to the EMBL Data Library, June 1987  
 A/Reference number: S17678  
 A/Accession: S17678  
 A/Molecule type: mRNA  
 A/Residues: 1-470, 'P', 472-682, 'G', 684-1043 <OBE>  
 A/Cross-references: EMBL:X05562; NID:G30075; PIDN:CAA29076.1; PID:G30076  
 R/Poeschl, E.; Pollner, R.; Kuehn, K.  
 EMBO J. 7, 2687-2695, 1988  
 A/Title: The gene for the alpha1(IV) and alpha2(IV) chains of human basement membrane  
 A/Reference number: S02738; MUID:89030632; PMID:2846280  
 A/Accession: S16911  
 A/Status: translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-33 <POS>  
 A/Cross-references: EMBL:X12784; GB:M36963; NID:G30072; PIDN:CAA1275.1; PID:G30073  
 R/Soininen, R.; Houtari, M.; Hoslika, S.L.; Prockop, D.J.; Trygsvaen, K.  
 J. Biol. Chem. 263, 17217-17220, 1988  
 A/Title: The structural genes for alpha1 and alpha2 chains of human type IV collagen ar  
 A/Reference number: A92690; MUID:89034231; PMID:3182844  
 A/Accession: B32117  
 A/Molecule type: DNA  
 A/Residues: 1-33 <SO11>  
 A/Cross-references: EMBL:J04217; EMBL:J05039; NID:G180759; PIDN:AAA53097.1; PID:G553233  
 R/Soininen, R.; Houtari, M.; Ganguly, A.; Prockop, D.J.; Trygsvaen, K.  
 J. Biol. Chem. 264, 13565-13571, 1989  
 A/Title: Structural organization of the gene for the alpha-1 chain of human type IV col  
 A/Reference number: S16877; MUID:89340433; PMID:2701944  
 A/Accession: S16877  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-33 <SO12>  
 A/Cross-references: EMBL:J04217; NID:G180759; PIDN:AAA53097.1; PID:G553233  
 A/Note: this sequence was submitted to the EMBL Data Library, October 1988  
 R/Siebold, B.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutmann, R.; Kuehn, K.  
 Eur. J. Biochem. 168, 569-575, 1987  
 A/Title: Construction of a model for the aggregation and cross-linking region (7S domai  
 is region.  
 A/Reference number: S00165; MUID:88029476; PMID:3117548  
 A/Accession: S00165  
 A/Molecule type: Protein  
 A/Residues: 37-247 <SE1>  
 A/Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-GI  
 R/Ehle, J.A.; Goldik, R.; Mann, K.; Kuehn, K.  
 EMBO J. 12, 4795-4802, 1993  
 A/Title: The alpha-1-beta-1 integrin recognition site of the basement membrane collagen  
 A/Reference number: S39614; MUID:94038963; PMID:8223488  
 A/Accession: S39615  
 A/Molecule type: Protein  
 A/Residues: 407-570 <EBL>  
 R/MacWright, R.S.; Benson, V.A.; Lovello, K.T.; van der Rest, M.; Fietzek, P.P.  
 Biochemistry 22, 4940-4948, 1983  
 A/Title: Isolation and characterization of pepsin-solubilized human basement membrane (I  
 A/Reference number: S16910; MUID:84053346; PMID:6416291  
 A/Accession: S16912  
 A/Molecule type: Protein  
 A/Residues: 490-492, 'X', 494-496, 675-677, 'G', 679-680, 'G', 682, 684-685, 'P' <MAC>  
 A/Experimental source: placenta



pol polyprotein - simian immunodeficiency virus SYIVagm (isolate SABD37) (fragment)  
 C:Species: simian immunodeficiency virus SYIVagm  
 A:Variety: isolate SABD37  
 C:Date: 25-Dec-1994 #sequence\_revision 14-Feb-1997 #text\_change 26-Aug-1999  
 C:Accession: S46354  
 R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;  
 EMOB J. 13, 2935-2947, 1994  
 A:Title: Mosaic genome structure of simian immunodeficiency virus from West African gree  
 A:Reference number: S46354; MUID:94298785; PMID:8026477  
 A:Accession: S46354  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-217 <JIN>  
 A:Cross-references: EMBL:U04018; NID:G466250; PIDN:AAA21512.1; PID:G466251  
 A:Experimental source: isolate SABD37; babaeus monkey  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993  
 C:Genetics:  
 A:Gene: pol  
 C:Superfamily: pol polyprotein  
 C:Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 217;  
 Best Local Similarity 75.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 EGPTLRQWLAR 9  
 :|||  
 Db 86 DGPRLRQW 93

RESULT 38  
 140327  
 baf protein - Bordetella pertussis  
 C:Species: Bordetella pertussis  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
 C:Accession: 140327; S70669  
 R:Deshaizer, D.; Wood, G.B.; Friedman, R.L.  
 J. Bacteriol. 177, 3801-3807, 1995  
 A:Title: Identification of a Bordetella pertussis regulatory factor required for transcr  
 A:Reference number: 140327; MUID:95325323; PMID:7601846  
 A:Accession: 140327  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-267 <RES>  
 A:Cross-references: UNIPROT:Q45338; EMBL:U12020; NID:G687228; PIDN:AAV5361.1; PID:G6872  
 R:Allen, A.; Maskell, D.  
 Mol. Microbiol. 19, 37-52, 1996  
 A:Title: The identification, cloning and mutagenesis of a genetic locus required for lip  
 A:Reference number: S70669; MUID:96419162; PMID:8821935  
 A:Accession: S70669  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 239-267 <ALL>  
 A:Cross-references: EMBL:X90711; NID:G992967; PIDN:CAA62242.1; PID:G992968  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995  
 C:Genetics:  
 A:Gene: baf

Query Match 53.4%; Score 39; DB 2; Length 267;  
 Best Local Similarity 63.6%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 3 GPTLRQWLAR 13  
 :|||  
 Db 195 GAIVRQWLAR 205

RESULT 39  
 B88445  
 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me  
 C:Species: Leishmania mexicana  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: B88445; S25142

R:Hannafert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.  
 Mol. Biochem. Parasitol. 55, 115-126, 1992  
 A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate  
 A:Reference number: A48445; MUID:93063042; PMID:1435864  
 A:Accession: B48445  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-331 <HAN>  
 A:Cross-references: UNIPROT:Q01556; EMBL:X65220; NID:G9552; PIDN:CAA46323.1; PID:G9553  
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase  
 C:Keywords: oxidoreductase

Query Match 53.4%; Score 39; DB 2; Length 331;  
 Best Local Similarity 38.5%; Pred. No. 57;  
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 IEPTLRQWLAR 13  
 :|||  
 Db 185 VDPSLRQWLAR 197

RESULT 40  
 A72514  
 hypothetical protein APE2086 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: A72514  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Cremonarchaeon, Aeropy  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: A72514  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-331 <KAW>  
 A:Cross-references: UNIPROT:Q9YA52; DDBJ:AP000063; NID:G5105654; PIDN:BAAB1097.1; PID:G  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2086  
 C:Superfamily: conserved hypothetical protein M01157

Query Match 53.4%; Score 39; DB 2; Length 331;  
 Best Local Similarity 50.0%; Pred. No. 57;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 GPTLRQWLARA 14  
 :|||  
 Db 306 GPVVRQWLARA 317

RESULT 41  
 C87021  
 serine-threonine protein kinase [imported] - Mycobacterium leprae  
 C:Species: Mycobacterium leprae  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: C87021  
 R:Coile, S.T.; Sigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: C87021  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-400 <STO>  
 A:Cross-references: UNIPROT:O69568; GB:AL450380; NID:G13092968; PIDN:CAC31278.1; GSPDB:  
 C:Genetics:  
 A:Gene: ML0897  
 C:Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific protein ki



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 15:48:37 ; Search time 52.0719 Seconds  
(without alignment)  
137.677 Million cell updates/sec

Title: US-10-083-768-13  
Perfect score: 73  
Sequence: 1 IEGLTLRQMLARA 14

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	68.5	302	2	Q742B3 mycobacteri
2	49.5	67.8	333	1	CBBR_XANFL
3	49	67.1	319	2	Q9RKM5 streptomyce
4	48	65.8	607	2	Q9L8D4 polyanthi
5	48	65.8	941	2	Q8QUC6
6	47	64.4	296	2	Q8ED06
7	46	64.4	296	2	Q8ZGS7
8	46	63.0	306	2	Q7D906 mycobacteri
9	46	63.0	306	2	O05576 mycobacteri
10	46	63.0	306	2	Q7U0W3
11	46	63.0	580	2	Q89RH2
12	45	61.6	245	2	O86272
13	45	61.6	249	2	O82989
14	45	61.6	278	2	Q9XDV0 erythroba
15	45	61.6	421	2	Q7W1X1
16	45	61.6	421	2	Q7W0U8
17	45	61.6	756	2	Q885P2
18	44	60.3	207	1	G1S3_CAEEL
19	44	60.3	536	1	ENTE_ECO57
20	44	60.3	536	1	ENTE_ECOLI
21	44	60.3	536	2	Q83M10
22	44	60.3	760	2	O8LMK9
23	43	58.9	285	1	KSGA_TREPA
24	43	58.9	297	2	Q7U0E4
25	43	58.9	306	2	O8ED00
26	43	58.9	354	2	O8ZYT5
27	43	58.9	377	2	O82PX5
28	43	58.9	683	2	O83436
29	43	58.9	754	2	O95Y82
30	42	57.5	91	2	O8Y0T5
31	42	57.5	126	2	O8N9N4

32	42	57.5	252	2	O8XP09	O8XP09 ralsstonia s
33	42	57.5	313	2	P04333	P04333 chimpanzee
34	42	57.5	325	2	O855N9	O855N9 mycobacteri
35	42	57.5	326	2	P95613	P95613 rhizobium g
36	42	57.5	375	2	Q7XP6	Q7XP6 oryza sativ
37	42	57.5	450	2	Q9SLB9	Q9SLB9 arabidopsis
38	42	57.5	586	2	O9N6P9	O9N6P9 leishmania
39	42	57.5	1019	1	POL_SIVS4	P12502 simian immu
40	42	57.5	1019	2	P89154	P89154 chimpanzee
41	42	57.5	1019	2	O7ZBR5	O7ZBR5 chimpanzee
42	42	57.5	1019	2	O7ZBR7	O7ZBR7 chimpanzee
43	41.5	56.8	410	2	O8P9L5	O8P9L5 xanthomonas
44	41.5	56.8	427	2	O8PLB2	O8PLB2 xanthomonas
45	41	56.2	75	2	O98A11	O98A11 rhizobium l
46	41	56.2	130	2	O6PBR0	O6PBR0 brachydanio
47	41	56.2	137	2	O9S720	O9S720 manihot pru
48	41	56.2	137	2	O9S721	O9S721 manihot esc
49	41	56.2	137	2	O9SPU6	O9SPU6 manihot esc
50	41	56.2	153	2	O80ZRO	O80ZRO mus musculu
51	41	56.2	173	2	O8GZB8	O8GZB8 brassica ra
52	41	56.2	195	1	ACPT_ECO57	O8X5U4 escherichia
53	41	56.2	195	1	ACPT_ECOLI	P37623 escherichia
54	41	56.2	195	1	O6N108	O6N108 rhodospendo
55	41	56.2	195	2	O6FCN3	O6FCN3 escherichia
56	41	56.2	207	2	O6PKJ2	O6PKJ2 orobanche m
57	41	56.2	209	2	O6N1X5	O6N1X5 rhodospendo
58	41	56.2	219	2	O8H6A7	O8H6A7 oryza sativ
59	41	56.2	223	2	O7XAP7	O7XAP7 houttuynia
60	41	56.2	224	2	O9F3Q7	O9F3Q7 streptomyce
61	41	56.2	244	2	O9R7K1	O9R7K1 erythroba
62	41	56.2	245	2	O82987	O82987 erythroba
63	41	56.2	245	2	O82991	O82991 erythroba
64	41	56.2	249	2	O9A535	O9A535 caulobacter
65	41	56.2	266	2	O81D70	O81D70 bacillus ce
66	41	56.2	268	2	O7Y082	O7Y082 lycopersico
67	41	56.2	294	2	O04891	O04891 lycopersico
68	41	56.2	306	2	O925G1	O925G1 mycobacteri
69	41	56.2	308	1	XRCR_CORGL	O8mnz9 corynebacte
70	41	56.2	313	2	O8WPF0	O8WPF0 capsicum an
71	41	56.2	313	2	G3PC_TOBAC	P09094 nicotiana t
72	41	56.2	326	1	G3PC_PETCR	P26519 petrobacteroi
73	41	56.2	336	2	O89ZV8	O89ZV8 bacteroides
74	41	56.2	337	1	G3PC_MESCR	P17878 mesembryant
75	41	56.2	337	2	G3PC_SINHL	P04796 sinapis alb
76	41	56.2	337	2	O6K5G8	O6K5G8 oryza sativ
77	41	56.2	338	1	G3PC_ARATH	P28588 arabidopsis
78	41	56.2	338	1	G3PC_DIACA	P34921 dianthus ca
79	41	56.2	338	2	O8LAK0	O8LAK0 arabidopsis
80	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
81	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
82	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
83	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
84	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
85	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
86	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
87	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
88	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
89	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
90	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
91	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
92	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
93	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
94	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
95	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
96	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
97	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
98	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
99	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub
100	41	56.2	338	2	O8LAK0	O8LAK0 solanum tub

## ALIGNMENTS



```

RESULT 1
Q742B3 PRELIMINARY; PRT; 302 AA.
AC 0742B3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Calu.
GN Name=gall; OrderedLocNames=MAP0924;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amosin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017230; AA03241.1;
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP_transferase; 1.
DR Complete proteome.
SQ SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;

Query Match 68.5%; Score 50; DB 2; Length 302;
Best Local Similarity 81.8%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAR 13
Db 286 GPDLRQWLVAR 296

RESULT 2
CBBR_XANFL STANDARD; PRT; 333 AA.
ID CBBR_XANFL
AC P2535;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE HTH-type transcriptional regulator cbbR (RubisCO operon
DE transcriptional regulator)
GN Name=cbbR; Synonyms=cfko;
OS Xanthobacter flavus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Hyphomicrobiaceae; Xanthobacter.
OX NCBI_TaxID=281;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H4-14;
RA MEDLINE=94012468; PubMed=8407781;
RA van den Bergen E., Dijkhuizen L., Meijer W.G.;
RT "CbbR, a LysR-type transcriptional activator, is required for
RT expression of the autotrophic CO2 fixation enzymes of Xanthobacter
RT flavus."
RT J. Bacteriol. 175:6097-6104(1993).
RN [2]
RP SEQUENCE OF 1-150 FROM N.A.
RC STRAIN=H4-14;
RA MEDLINE=9112133; PubMed=1900916;
RA Meijer W.G., Arberg A.C., Enequist H.G., Terpstra P., Lidstrom M.E.,
RA Dijkhuizen L.;
RT "Identification and organization of carbon dioxide fixation genes in
RT Xanthobacter flavus H4-14."
RC Mol. Gen. Genet. 225:320-330(1991).
CC -1- FUNCTION: Transcriptional activator for the cbb operon (cbbLSXFP)
CC for RubisCO and other Calvin cycle genes. Binds specifically to
CC two binding sites in the cbbP-cbbL intergenic region.
CC -1- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22705; CAAB0406.1;
DR EMBL; X17252; -; NOT_ANNOTATED_CDS.
DR PIR; A36925; A36925.
DR InterPro; IPR000847; HTH_LYER.
DR InterPro; IPR005119; LysR_subst.
DR InterPro; IPR009058; Wnt_hlx_DNA_bnd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYER; 1.
KM Activator; DNA-binding; Transcription regulation.
FT DOMAIN 5 62 HTH_LysR-type.
FT DNA_BIND 22 41 H-T-H motif (By similarity).
SQ SEQUENCE 333 AA; 36003 MW; 9B375B4FB2D1E873 CRC64;

Query Match 67.8%; Score 49.5; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 IEQ-PTLRQWLARA 14
Db 264 VEGLPVVRQWLVARA 278

RESULT 3
Q9RKM5 PRELIMINARY; PRT; 319 AA.
ID Q9RKM5
AC Q9RKM5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MerR family transcriptional regulator.
GN ORFNames=SCD17.06c;
OS Streptomyces coelicolor.
OC Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/41741a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Baleman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Hwang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Ruter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RT Nature 417:141-147(2002).
CC -1- SIMILARITY: Contains 1 HTH merR-type DNA-binding domain.
CC EMBL; AL939118; CAB56383.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR InterPro; IPR009061; Putativ_DNA_bind.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MER_R; 1.
DR PROSITE; PS50937; HTH_MER_R; 1.
KM Complete proteome; DNA-binding.

```



SEQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;

Query Match 67.1%; Score 49; DB 2; Length 319;

Best Local Similarity 66.7%; Pred. No. 4;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAR 13

DB 258 DGPRLRWLAQR 269

RESULT 4

Q9L8D4 PRELIMINARY; PRT; 607 AA.

AC Q9L8D4; (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

OS Hypothetical protein (fragment).

OC Polyanthum celluloseum (Sorangium cellulosum).

OC Bacteria; Proteobacteria; Delaproteobacteria; Myxococcales;

OC Sorangineae; Polyangifacae; Polyangium.

OC NCBI\_TaxID=56;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=So ce90;

RX MEDLINE=20130945; PubMed=10662695; DOI=10.1016/S1074-5521(00)00075-2;

RA Mojar I., Schupp T., Ono M., Zirkle R.E., Milnamow M.,

RA Novak-Thompson B., Engel N., Toupet C., Strattmann U., Cyr D.D.,

RA Goralach J., Mayo J.M., Hu A., Goff S., Schmidt J., Lion J.M.;

RT "The biosynthetic gene cluster for the microtubule-stabilizing agents

RT epothilones A and B from Sorangium cellulosum So ce90.";

RL Chem. Biol. 7:97-109 (2000).

DR EMBL; AF210843; AAF26904.1; -.

KM Hypothetical protein.

FT NOK TR 1

SEQ SEQUENCE 607 AA; 66326 MW; F113CA299B25048E CRC64;

Query Match 65.8%; Score 48; DB 2; Length 607;

Best Local Similarity 61.5%; Pred. No. 11;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAR 13

DB 96 VDGPAVYRWLAAR 108

RESULT 5

Q8QUU6 PRELIMINARY; PRT; 941 AA.

AC Q8QUU6; (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

OS Infectious spleen and kidney necrosis virus.

OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae;

OC unclassified Iridoviridae.

OC NCBI\_TaxID=180170;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;

RA He J.G., Deng M., Wang S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,

RA Chan S.M.;

RT "Complete genome analysis of the mandarin fish infectious spleen and

RT kidney necrosis iridovirus.";

RL Virology 291:126-139 (2001).

DR EMBL; AF31960; AAL98838.1; -.

SEQ SEQUENCE 941 AA; 106703 MW; EB663998C7F6CE83 CRC64;

Query Match 65.8%; Score 48; DB 2; Length 941;

Best Local Similarity 50.0%; Pred. No. 18;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEPTLRQWLARA 14

DB 581 VQPTLRQWLICSTRA 594

RESULT 6

Q66D06 PRELIMINARY; PRT; 296 AA.

AC Q66D06; (TREMBLrel. 28, Created)

DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)

DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

OS Putative drug/metabolite (DME family) efflux pump precursor.

OR Names=YPM1243;

OS Yersinia pseudotuberculosis IP 32953.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OC NCBI\_TaxID=273123;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IP 32953;

RX PubMed=15358858;

RA Chain P.S.G., Carniel E., Larimer F.W., Iamerdin J., Stoutland P.O.,

RA Regala M.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,

RA Brubaker R.R., Fowler J., Hinebusch B.J., Marceau M., Medigue C.,

RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,

RA Derdier A., Hauser L.J., Garcia E.;

RT "Insights into the genome evolution of Yersinia pestis through whole

RT genome comparison with Yersinia pseudotuberculosis.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831 (2004).

DR EMBL; BX96398; CAH20483.1; -.

DR InterPro; IPR00620; DUF6.

DR Pfam; PF00892; DUF6; 2.

DR Signal.

FT SIGNAL.

SEQ SEQUENCE 296 AA; 31407 MW; 4D3E486D32DBAC11 CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;

Best Local Similarity 81.8%; Pred. No. 8.2;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PTLRQWLARA 14

DB 66 PTLRQWLARA 76

RESULT 7

Q8ZGS7 PRELIMINARY; PRT; 296 AA.

AC Q8ZGS7; Q74WE0; Q7CH89;

DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

DE Putative membrane protein (Putative transmembrane protein).

GN Name=thata; OrderedLocustNames=YPO934, YPO1203, Y285;

OS Yersinia pestis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OC NCBI\_TaxID=632;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=CO-92 / Bover Orientalis;

RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;

RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,

RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,

RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Fellwell T., Hamlin N., Holtroyd S., Jags K., Kariyeh A.V.,

RA Leather S., Moul S., Oyston P.C.F., Quail M.A., Rutherford K.M.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;

RT "Genome sequence of Yersinia pestis, the causative agent of plague.";

RL Nature 413:523-527 (2001).

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,
RA Ferns N.T., Rose D.J., Meun B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ414147; CAC90042.1; -.
DR EMBL: AE013900; AAM86536.1; -.
DR EMBL: AE017130; AAS61189.1; -.
DR PIR: AG01477; AG0147.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR Pfam: PF00892; DUF6; 2.
DR Complete proteome; Transmembrane.
SQ SEQUENCE 296 AA; 31378 MW; 45947413DCD54CF6 CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred. No. 8.2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PTLRQWLAR 14
Db 66 PTLRQWLAR 76

RESULT 8
Q79006 PRELIMINARY; PRT; 306 AA.
AC Q79006;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DR UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9).
GN Name=gall; OrderedLocNames=MT1022;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.C., Umayam L.A., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL: AE000516; AAK45263.1; -.
DR TIGR: MT1022; -.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0003983; F:UTP-glucose-1-phosphate uridylyltransferase. . .; IEA.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR InterPro: IPR005835; NTP transferase.
DR Pfam: PF00483; NTP_transferase; 1.
DR Nucleotidyltransferase; Transferase.

```

```

SQ SEQUENCE 306 AA; 32406 MW; 880D3BB86CBA3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAR 13
Db 290 GPTLRQWLAR 300

RESULT 9
O05576 PRELIMINARY; PRT; 306 AA.
AC O05576;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DR PROBABLE UTP-glucose-1-phosphate uridylyltransferase GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE) (EC
DE 2.7.7.9).
GN Name=gall; OrderedLocNames=Rv0933;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RX Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RX Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RX Connor R., Davies K.M., Devlin K., Feldwell T., Gentles S., Hamlin N.,
RX Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RX Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RX Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RX Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
DR EMBL: BX842575; CAB08153.1; -.
DR PIR: D70601; D70601.
DR Tuberculosis; Rv0933; -.
DR GO: GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR InterPro: IPR005835; NTP transferase.
DR Pfam: PF00483; NTP_transferase; 1.
DR Complete proteome.
SQ SEQUENCE 306 AA; 32378 MW; 24C2387443B0A3E8 CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAR 13
Db 290 GPTLRQWLAR 300

RESULT 10
Q7U0W3 PRELIMINARY; PRT; 306 AA.
AC Q7U0W3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR PROBABLE UTP-glucose-1-phosphate uridylyltransferase GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE) (EC
DE 2.7.7.9).
GN Name=gall; OrderedLocNames=Mb1020;

```

```

OS Mycobacterium bovis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Bigmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Gordin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hwinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO:0016740; F:transferase activity; IEA.
DR GO:0009058; P:biosynthesis; IEA.
DR InterPro: IPR005835; NTP transferase.
DR Pfam: PF00483; NTP transferase; 1.
KM Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32406 MW; 880D3BB8CB0A3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GPTLRQWLAR 13
Db 290 GPDLRHMLVAR 300

RESULT 11
O89RH2 PRELIMINARY; PRT; 580 AA.
ID O89RH2
AC O89RH2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE B112800 protein.
GN OrderedLocustNames=b112800;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobiium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Itiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL: AP005945; BAC48065.1; -.
DR GO:0005524; F:ATP binding; IEA.
DR GO:0003824; F:catalytic activity; IEA.
DR GO:0004672; F:protein kinase activity; IEA.
DR GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR001932; P2C-like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00481; P2C; 1.
DR ProDom: PD00001; Prot_kinase; 1.
DR SMART: SM00332; P2C; 1.
DR SMART: SM00331; P2C_SIG; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 580 AA; 64916 MW; 6AD3A06BE6FAE143B CRC64;

```

```

Query Match 63.0%; Score 46; DB 2; Length 580;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 IEGPTLRQWL 10
Db 355 IEQGTLRQWL 364

RESULT 12
O62272 PRELIMINARY; PRT; 245 AA.
ID O62272
AC O62272;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufl;
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMM14332; PubMed=11832943; DOI=10.1038/415630a;
RX MEDLINE=21822632;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL: AB010981; BAA25791.1; -.
DR HSPF: P02954; 10OV.
DR GO:0030077; C:light-harvesting complex (sensu Proteobact. . .; IEA.
DR GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO:0006118; P:electron transport; IEA.
DR GO:0019684; P:photosynthesis, light reaction; IEA.
DR InterPro: IPR005871; Photo_L.
DR InterPro: IPR00484; Photo_RC.
DR Pfam: PF00124; Photo_RC; 1.
DR PRINTS: PR00256; REACTCENTRE.
DR TIGRFAMs: TIGR01157; pufl; 1.
DR PROSITE: PS00244; REACTION_CENTER; 1.
FT NON TER 1
SQ SEQUENCE 245 AA; 27214 MW; 52B268733E199A8D CRC64;

Query Match 61.6%; Score 45; DB 2; Length 245;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 IEGPTLRQWL 10
Db 26 IEQTLNPWL 35

RESULT 13
O82989 PRELIMINARY; PRT; 249 AA.
ID O82989
AC O82989;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufl;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBIC3019; PubMed=11832943; DOI=10.1038/415630a;
RX MEDLINE=21822632;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,

```

```

RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL: AB015708; BAA32995.1; -.
DR HSSP: P02954; 1YST.
DR GO: GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO: GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0019684; P:photosynthesis, light reaction; IEA.
DR InterPro: IPR005871; Photo_L.
DR InterPro: IPR000484; Photo_RC.
DR Pfam: PF00124; Photo_RC.1_RC.
DR PRINTS: PR00256; REACTCENTRE.
DR TIGRFAMs: TIGR01157; pufL.1.
DR PROSITE: PS00244; REACTION_CENTER; 1.
FT NON TER 1
SQ SEQUENCE 249 AA; 27702 MW; 4D68BDC82B7166AD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 249;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IEGETLAWL 10
Db 26 IEGETLAWL 35

RESULT 14
O9XDV0 PRELIMINARY; PRT; 278 AA.
AC O9XDV0,
ID 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit.
GN Name-pufL;
OS Erythrobacter sp. MBIC3960.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_Taxid=94771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBIC3960;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Bja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL: AB027515; BAA78672.1; -.
DR HSSP: P02954; 1YST.
DR GO: GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO: GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0019684; P:photosynthesis, light reaction; IEA.
DR InterPro: IPR005871; Photo_L.
DR InterPro: IPR000484; Photo_RC.
DR Pfam: PF00124; Photo_RC.1.
DR PRINTS: PR00256; REACTCENTRE.
DR TIGRFAMs: TIGR01157; pufL.1.
DR PROSITE: PS00244; REACTION_CENTER; 1.
SQ SEQUENCE 278 AA; 30735 MW; 0BE618844B3C54FB CRC64;

Query Match 61.6%; Score 45; DB 2; Length 278;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IEGETLAWL 10
Db 55 IEGETLAWL 64

RESULT 15
O7WIX1

```

```

ID O7WIX1 PRELIMINARY; PRT; 421 AA.
AC O7WIX1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedlocusNames=BP0223;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_Taxid=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL: BX640423; CAE39964.1; -.
DR GO: GO:0016874; F:ligase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 2.
KW Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45579 MW; 13D6606AFLPDEC21 CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 4 PTUROWLAAR 13
Db 221 PSURDWLAAR 230

RESULT 16
O7WQUB PRELIMINARY; PRT; 421 AA.
ID O7WQUB;
AC O7WQUB;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedlocusNames=BB0227;
OS Bordetella bronchiseptica.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_Taxid=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,

```

```

RT Bordetella parapertussis and Bordetella bronchiseptica.",
RL Mac. Gene. 35:32-40(2003).
DR EMBL; BX640437; CAE30725.1; -.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
KW Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45558 MW; A6CDBC9C731A49C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PTLROWLAAR 13
   |||||
Db 221 PSLRDWLAAR 230

RESULT 17
Q885P2 PRELIMINARY; PRT; 756 AA.
AC Q885P2:
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Dimeric lysozyme reductase.
GN Ordered locus names=PSPT01789;
OS Pseudomonas syringae (pv. tomat).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buehl C.R., Joariz V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Davsberry S.C., Binkac L.M., Beaman M.J., Halt D.H.,
RA Nelson W.C., Davidsen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Pedorova N.B., Tran B., Russell D., Berry K.J.,
RA Uteerbeck T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartimour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomat DC3000".
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016862; AAO55309.1; -.
DR HSSP; Q57366; 1E01.
DR TIGR; PSP01789; -.
DR GO; GO:0030151; F:molybdenum ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR009010; Asp-decarb fold.
DR InterPro; IPR006557; Mol_dinuc_bind.
DR Pfam; PF01568; Molybdop_binding; 1.
KW Complete proteome.
SQ SEQUENCE 756 AA; 83189 MW; 31E9614DE2B22B2C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 756;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPLROWLAAR 13
   |||||
Db 222 LAGPTHOWLAAR 234

RESULT 18
GTS3_CABEL STANDARD; PRT; 207 AA.
AC O16116; Q21357;
DT 10-OCT-2003 (Rel. 42, Created)

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glutathione S-transferase 3 (EC 2.5.1.18) (GST class-sigma) (ceGSr3).
GN Name=gst-3; ORFNames=K08F4.11;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Tawe W.N., Bachrach M.-L., Walter R.D., Henkle-Duehn K.;
RT "Paracat mediates differential gene expression in C. elegans.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP REVISIONS.
RA Durbin R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC exogenous and endogenous hydrophobic electrophiles (by
CC similarity).
CC -1- SIMILARITY: Belongs to the GST superfamily. Sigma family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF010241; AAB65419.1; -.
DR EMBL; Z68879; CA93088.2; -.
DR PIR; T37464; T37464.
DR HSSP; P24472; 1GUK.
DR WormBase; WBGene00001751; gst-3.
DR WormPep; K08F4.11; CE25050.
DR GO; GO:0004364; F:glutathione transferase activity; ISS.
DR InterPro; IPR010987; GST_C-like.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
KW Transferase.
SQ SEQUENCE 207 AA; 23735 MW; 72545319FCFCEBDA CRC64;

Query Match 60.3%; Score 44; DB 1; Length 207;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPLROWLAAR 13
   |||||
Db 190 IETPLKWLAKR 202

RESULT 19
ENTE_ECOS7 STANDARD; PRT; 536 AA.
AC Q8XBV3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Enterobacteriaceae synthease component B (Enterobacteriaceae synthease B)
DE [includes: 2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)]

```

```

DE (Dihydroxybenzoic acid-activating enzyme); S-
DE dihydroxybenzoyltransferase [EC 2.3.1.-)].
CN Name=entB; OrderedLocustNames=20736, EC06633;
OS Escherichia coli O157:H7.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11205551; DOI=10.1038/35054089;
RA Perma N.T., Plunkett G., Ilt, Burtland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Mackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta B.T., Potamoustis K.,
RA Apodaca J., Anantharaman T.S., Iln J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RC MEDLINE=2115633; PubMed=11258796;
RX Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasekawa C., Ogasaawara N., Yasunaga T.,
RA Kumura S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
CC -I- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
CC benzoate (2,3-DHB), via ATP-dependent PPI exchange reactions, to
CC the acyladenylate. Then, catalyzes the acylation of holo-entB with
CC 2,3-DHB adenylate, preparing that molecule for amide bond
CC formation with L-serine (By similarity).
CC -I- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
CC (2,3-dihydroxybenzoyl)-adenylate.
CC -I- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)-adenylate + holo-entB =
CC adenosine 5'-monophosphate + acyl-holo-entB.
CC -I- PATHWAY: siderophore biosynthesis; enterobactin biosynthesis.
CC -I- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
CC complex called enterobactin synthase (By similarity).
CC -I- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family. EntB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005239; AAG54929.1; -
CC EMBL; AP002552; BAB34056.1; -
CC PIR; A99708; A99708.
CC HSSP; P40871; 1MD9.
CC InterPro; IPR00873; AMP-bind.
CC Pfam; PF00501; AMP-binding; 1.
CC DR TIGRFAMS; TIGR01733; AA-adenyl-dom; 1.
CC DR TIGRFAMS; TIGR01923; menE; 1.
CC PROSITE; PS00455; AMP BINDING; 1.
CC Acyltransferase; Complete proteome; Enterobactin biosynthesis;
CC Iron transport; ligase; Multifunctional enzyme; Transferase;
CC Transport.
CC SEQUENCE 536 AA; 59040 MW; ABC8EDB320940A5 CRC64;
SQ

```

```

RESULT 20
DB 521 VDRKQURQWLASRA 534

ID      ENTER_ECOLI      STANDARD;      PRT;      536 AA.
AC      P10378; P15049; P77773;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      25-JAN-2005 (Rel. 46, Last annotation update)
DE      Enterobactin synthetase component E (Enterobactin synthase E)
DE      [includes: 2,3-dihydroxybenzoate-AMP lyase (EC 2.7.7.56)
DE      2,3-dihydroxybenzoic acid-activating enzyme); S-
DE      dihydroxybenzoyltransferase (EC 2.3.1.-)].
DE      Name=entB; OrderedLocustName=bd0594;
DE      Escherichia coli.
DE      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
DE      Enterobacteriaceae; Escherichia.
CN      NCBI_TaxId=562;
NN
NN      SEQUENCE FROM N.A.
NN      STRAIN=K12;
NN      MEDLINE=69280355; PubMed=2525505; DOI=10.1016/0378-1097(89)90450-3;
NN      Straub J.F., Elkins M.F., Earhart C.F.;
NN      "Nucleotide sequence of the Escherichia coli ente gene.";
NN      FEBS Microbiol. Lett. 50:15-19(1989).
NN      [2]
NN      SEQUENCE FROM N.A.
NN      STRAIN=K12 / M61655;
NN      MEDLINE=97446617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
NN      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
NN      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
NN      Gregor Y., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
NN      Mau B., Shao Y.;
NN      "The complete genome sequence of Escherichia coli K-12.";
NN      Science 277:1453-1474(1997).
NN      [3]
NN      SEQUENCE FROM N.A.
NN      Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
NN      Federpiel N., Hyman R., Kallman S., Komp C., Kurd O., Lew H., Lin D.,
NN      Namatb A., Oefner P., Roberts D., Schramm S., Davis R.W.;
NN      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
NN      [4]
NN      SEQUENCE OF 1-8 FROM N.A.
NN      STRAIN=K12;
NN      MEDLINE=90236256; PubMed=2110093; DOI=10.1016/0378-1097(90)90120-F;
NN      Elkins M.F., Earhart C.F.;
NN      "Opacity factor from group A streptococci is an apoproteinase.";
NN      FEBS Microbiol. Lett. 56:35-40(1988).
NN      [5]
NN      SEQUENCE OF 393-546 FROM N.A.
NN      MEDLINE=89123155; PubMed=2511622;
NN      Liu J., Duncan K., Walsh C.T.;
NN      "Nucleotide sequence of a cluster of Escherichia coli enterobactin
NN      biosynthesis genes: identification of entB and purification of its
NN      product 2,3-dihydroxy-2,3-dihydroxybenzoate dehydrogenase.";
NN      J. Bacteriol. 171:791-798(1989).
NN      [6]
NN      FUNCTION.
NN      MEDLINE=97361959; PubMed=9214294; DOI=10.1021/bi970453p;
NN      Gearing A.M., Bradley K.A., Walsh C.T.;
NN      "Enterobactin biosynthesis in Escherichia coli: isochorismate lyase
NN      (Entb) is a bifunctional enzyme that is phosphoenolpyruvate-
NN      Entd and then acylated by Entf using ATP and 2,3-dihydroxybenzoate.";
NN      Biochemistry 36:8485-8503(1997).
NN      -1- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
NN      benzoate (2,3-DHB), via ATP-dependent PPI exchange reactions, to
NN      the acyladenylate, then, catalyzes the acylation of holo-entB with
NN      2,3-DHB adenylate, preparing that molecule for amide bond
NN      formation with L-serine.
NN      -1- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
NN      (2,3-dihydroxybenzoyl)-adenylate.
NN      -1- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)-adenylate + holo-entB =

```

```

CC      adenosine 5'-monophosphate + acyl-holo-entB.
CC      -1- PATHWAY: Siderophore biosynthesis; enterobactin biosynthesis.
CC      -1- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
CC      complex called enterobactin synthetase.
CC      -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC      family. EntB subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U00096; AAC73695.1; -
CC      EMBL; U82598; AAB40794.1; -
CC      EMBL; X15058; CA33158.1; -
CC      EMBL; M24148; AAA16101.1; -
CC      EMBL; M36700; AAA18492.1; -
CC      PIR; H64792; SYCEB.
CC      HSSP; P40871; IMD9.
CC      EcoBASE; EB0259; -.
CC      EcoGene; EG10263; entB.
CC      InterPro; IPR000873; AMP-bind.
CC      Pfam; PF00501; AMP-binding; 1.
CC      PRINTS; PR00154; AMPBINDING.
CC      TIGRFAMs; TIGR01733; AA-adenyl-dom; 1.
CC      TIGRFAMs; TIGR01923; menB; 1.
CC      PROSITE; PS00455; AMP BINDING; 1.
CC      Acyltransferase; Complete proteome; Enterobactin biosynthesis;
CC      Iron transport; Ligase; Multifunctional enzyme; Transferase;
CC      Transport.
CC      FT CONFLICT 369 AA; 378 DAEGNPLPQG -> ECRRKSTAR (in Ref. 1).
SQ      SEQUENCE 536 AA; 59112 MW; F818942DFD8DC99 CRC64;

Query Match      60.3%; Score 44; DB 1; Length 536;
Best Local Similarity 57.1%; Pred. No. 50;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEGPTLRQWLARA 14
Db      521 VDKKQLRQWLASRA 534

RESULT 21
O83M10      PRELIMINARY;      PRT;      536 AA.
AC      O83M10; Q7C2S3;
DT      01-JUN-2003 (TREMBlrel. 24, Created)
DT      01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT      25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE      2,3-dihydroxybenzoate-AMP ligase.
GN      Name=entB; OrderedLocNames=S0514, SFO508;
OS      Shigella flexneri.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Shigella.
OX      NCBI_TaxID=623;
ON      [1]
RA      Yu J.;
RA      "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RA      through comparison with genomes of Escherichia coli K12 and O157.",
RA      Nucleic Acids Res. 30:4432-4441 (2002).
RL      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=2457T;

```

```

RX      MEDLINE=22590274; PubMed=12704152;
RX      DOI=10.1128/JAI.71.5.2775-2786.2003;
RA      Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA      Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA      Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA      Schwartz D.C., Blattner F.R.;
RT      "Complete genome sequence and comparative genomics of Shigella
RT      flexneri serotype 2a strain 2457T.",
RL      Infect. Immun. 71:2775-2786 (2003).
CC      -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC      family.
CC      EMBL; AE015082; AAN42156.1; -
CC      EMBL; AE016979; AAP16028.1; -
CC      HSSP; P40871; IMD9.
CC      DR      GO; GO:0016874; F:ligase activity; IEA.
CC      DR      GO; GO:0008152; P:metabolism; IEA.
CC      DR      InterPro; IPR000873; AMP-bind.
CC      DR      Pfam; PF00501; AMP-binding; 1.
CC      DR      PRINTS; PR00154; AMPBINDING.
CC      DR      PROSITE; PS00455; AMP BINDING; UNKNOWN_1.
CC      KW      Ligase; Complete proteome.
SQ      SEQUENCE 536 AA; 58851 MW; ABABD6B8692ABD2 CRC64;

Query Match      60.3%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 50;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEGPTLRQWLARA 14
Db      521 VDKKQLRQWLASRA 534

RESULT 22
O8LMK9      PRELIMINARY;      PRT;      760 AA.
AC      O8LMK9;
DT      01-OCT-2002 (TREMBlrel. 22, Created)
DT      01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT      05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE      Putative gag/pol polyprotein (Putative gag-pol polypeptide).
GN      ORFNames=OSJNB003812.4;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzaceae; Oryza.
OX      NCBI_TaxID=39947;
ON      [1]
RA      Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA      Overton I.T., Taitlin T., Kim M.M., Beta J.J., Jin S.S.,
RA      Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Heitao J., Blunt S.,
RA      Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblum T.V.,
RA      Yang O.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA      White O., Salzberg S.L., Fraser C.M.;
RL      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Buell R.;
RL      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      "The Rice Chromosome 10 Sequencing Consortium;
RA      "In-depth view of structure, activity, and evolution of rice
RA      chromosome 10.",
RL      Science 300:1566-1569 (2003).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RA      Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AC105932; AAN04966.1; -
DR      EMBL; AE017067; AAP52546.1; -
DR      Gramene; O8LMK9; -.
DR      InterPro; IPR005162; Retrotrans_gag.

```



DR Pfam; PF03732; Retrotrans\_gag; 1.  
 KM Polyprotein.  
 SQ SEQUENCE 760 AA; 82020 MW; C51F91AA2EB32A28 CRC64;

Query Match 60.3%; Score 44; DB 2; Length 760;  
 Best Local Similarity 46.2%; Pred. No. 71;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPLRLQWLAAR 13  
 : ||||| : ||| :  
 Db 661 LHGPTLQHWMAVK 673

RESULT 23  
 KSGA TREPA STANDARD; PRT; 285 AA.

AC 083357;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N',  
 DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High  
 DE level kasugamycin resistance protein ksgA) (Kasugamycin  
 DE dimethyltransferase).  
 GN Name=ksgA; OrderedLocustNames=TP0337;  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Nichols;  
 RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,  
 RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,  
 RA Uterback T.R., McDonald L.A., Atlich P., Bowman C., Cotton M.D.,  
 RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,  
 RA Weidman J.F., Smith H.O., Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete";  
 RL Science 281:375-388(1998).

CC -1- FUNCTION: Specifically, dimethylates two adjacent adenosines in the  
 CC loop of a conserved hairpin near the 3' end of 16S rRNA in the 30S  
 CC particle. Its inactivation leads to kasugamycin resistance (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the RNA adenine N-6-methyltransferase  
 CC family. KsgA subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AB001213; AAC65323.1; -

DR PIR; G71337; G71337.  
 DR TIGR; TP0337; -  
 DR HAMAP; MF\_00607; -, 1.

DR InterPro; IPR001737; RNA\_A\_dimeth.  
 DR InterPro; IPR000051; SAM\_Bind.  
 DR Pfam; PF00398; Rnamd; 1.  
 DR SMART; SMO0650; RADC; 1.

DR TIGRFAMs; TIGR00755; ksgA; 1.  
 DR PROSITE; PS01131; RNA\_A\_DIMETH; FALSE\_NEG.

KM Antibiotic resistance; Complete proteome; Methyltransferase;  
 KM RNA processing; Transferase.  
 SQ SEQUENCE 285 AA; 32275 MW; 3AF0BCBE16B5DAF CRC64;

Query Match 58.9%; Score 43; DB 1; Length 285;  
 Best Local Similarity 64.3%; Pred. No. 39;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPLRLQWLAARA 14  
 ||||| : ||||| :  
 Db 98 IEGDVLQWMAAAA 111

RESULT 24  
 Q7UOE4 PRELIMINARY; PRT; 297 AA.

AC Q7UOE4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=RB6375;  
 OS Rhodospirillum rubrum.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,  
 RA Schlensker H., Amani R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).

DR EMBL; BX294144; CAD74759.1; -  
 DR InterPro; IPR000194; ATPase\_a/bcentre.  
 DR InterPro; IPR003169; GYF.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; UNKNOWN\_1.  
 DR PROSITE; PS50829; GYF; 1.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 297 AA; 31805 MW; 475F670F02C7859B CRC64;

Query Match 58.9%; Score 43; DB 2; Length 297;  
 Best Local Similarity 50.0%; Pred. No. 41;  
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGPLRLQWLA 11  
 : ||||| : ||| :  
 Db 176 DGPTMKWIS 185

RESULT 25  
 Q8EJ00 PRELIMINARY; PRT; 306 AA.

AC Q8EJ00;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Prophage Mus81, major head subunit, putative.  
 GN OrderedLocustNames=SO0675;  
 OS Shewanella oneidensis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Shewanellaceae; Shewanella.  
 OX NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=MS-1;  
 RX MEDLINE=2237686; PubMed=12368813; DOI=10.1038/nbt749;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Sehadri R., Ward N.L., Methe B.A.,  
 RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beaman M.J.,  
 RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,  
 RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,  
 RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Imyrain M.,  
 RA Lee K., Berry K.J., Lee C., Mueller J., Kouri H.M., Gill J.,  
 RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,  
 RA Venter J.C., Neilson K.H., Fraser C.M.;



```

RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.",
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015513; AANS3753.1; -.
DR TIGR; SC0675; -.
KM Complete proteome.
SQ SEQUENCE 306 AA; 34370 MW; F54CCA118AA288CB CRC64;

Query Match 58.9%; Score 43; DB 2; Length 306;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTLRQWLAAR 13
|||:|:|
Db 54 PTMEWTGAR 63

RESULT 26
ID 082YTS PRELIMINARY; PRT; 354 AA.
AC 082YTS;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein PAE0634.
GN OrderedLocustNames=PAE0634;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IN2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stettler K.O., Simon M.I.,
RA Miller J.H.,
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009776; ALU62908.1; -.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 354 AA; 38642 MW; C5799P975B972941 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 354;
Best Local Similarity 61.5%; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
|||:|:|
Db 84 IDRGLEQWLAAR 96

RESULT 27
ID 082PK5 PRELIMINARY; PRT; 377 AA.
AC 082PK5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SAV747;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinomycetia; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kituchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces

```

```

RT avermitilis: deducing the ability of producing secondary
RT metabolites.",
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AE001220; AAC65409.1; -.
DR TIGR; TP0421; -.
DR InterPro; IPR011044; Amine_DH_B_like.
DR InterPro; IPR002110; AMK.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR001440; TPR.
DR Pfam; PF01436; NHL; 5.
DR PRINTS; PR01415; ANKYRN.
DR PROSITE; PSS0005; TPR; 1.
DR PROSITE; PSS0293; TPR_REGION; 1.
KM Complete proteome.
SQ SEQUENCE 683 AA; 74518 MW; F91407FA7094AD1 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 95;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
|||:|:|
Db 89 IEQAALHMGGAAR 101

RESULT 28
ID 083436 PRELIMINARY; PRT; 683 AA.
AC 083436;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein TP0421.
GN OrderedLocustNames=TP0421;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=9833770; PubMed=965876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwin M.L., Hickey B.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Ueberlack T.R., McDonald L.A., Artlich P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.",
RL Science 281:375-388(1998).
DR EMBL; AE001220; AAC65409.1; -.
DR PIR; B71325; B71325.
DR TIGR; TP0421; -.
DR InterPro; IPR011044; Amine_DH_B_like.
DR InterPro; IPR002110; AMK.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR001440; TPR.
DR Pfam; PF01436; NHL; 5.
DR PRINTS; PR01415; ANKYRN.
DR PROSITE; PSS0005; TPR; 1.
DR PROSITE; PSS0293; TPR_REGION; 1.
KM Complete proteome.
SQ SEQUENCE 683 AA; 74518 MW; F91407FA7094AD1 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 95;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
|||:|:|
Db 89 IEQAALHMGGAAR 101

```

RESULT 29

095Y82 PRELIMINARY; PRT; 754 AA.

AC 095Y82;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypoetical protein Y119C1B.5.

GN Name=Y119C1B.5; ORFNames=Y119C1B.5;

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RG Wormbase Consortium;

RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Jones K., Murray J., Graves T.;

RT "The sequence of *C. elegans* cosmid Y119C1B.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RN [7]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

RN [8]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RN [9]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Wilson R.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

RN [10]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Wormbase Consortium;

RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

CC -1. SIMILARITY: Contains 1 RING-type zinc finger.

DR EMBL; AC006712; AAK39324.1; -.

DR Wormbase; WBGene00022471; Y119C1B.5.

DR Wormpep; Y119C1B.5; CE27234.

DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0016567; P:protein ubiquitination; IEA.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS50089; ZF\_RING\_2; 1.

KW Hypoetical protein; Metal-binding; Zinc; Zinc-finger.

SQ SEQUENCE 754 AA; 85323 MW; 41BA9297FA3BF05 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 754;

Best Local Similarity 63.6%; Pred. No. 1.1e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAR 13

Db 626 GPCLRKWLAVK 636

RESULT 30

08Y015 PRELIMINARY; PRT; 91 AA.

AC 08Y015;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypoetical protein RSC1059.

GN Name=RS04149; OrderedLocNames=RS041059;

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Burkholderiaceae; Ralstonia.

OX NCBI\_TaxID=505;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GW11000;

RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;

RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S., Ariat M., Billault A., Brottier P., Camus J.C., Catolico L., Chandler M., Cholene N., Claudel-Renard C., Cunnac S., Demange N., Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T., Siguler P., Thebaud P., Whalen M., Wincker P., Levy M., Weissbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";

RL Nature 415:497-502(2002).

DR EMBL; AL646062; CAD14761.1; -.

KW Complete proteome.

SQ SEQUENCE 91 AA; 10321 MW; 2B4DFEB37A528AD CRC64;

Query Match 57.5%; Score 42; DB 2; Length 91;

Best Local Similarity 46.2%; Pred. No. 18;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAR 13

Db 75 LDGPVQAWMLAQ 87

RESULT 31

08N9N4 PRELIMINARY; PRT; 126 AA.

AC 08N9N4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypoetical protein FLJ36840.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC PubMed=14702039; DOI=10.1038/ng1285;

RX Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,

RA Makamatsu A., Hayaishi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahata K.,  
RA Murakami K., Saito T., Iwayanagi T., Wagatsuna M., Shiratori A.,  
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
RA Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,  
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T., Kusano J.,  
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
RA Masehino K., Yuki H., Oshima A., Sasaki N., Aochi S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Senba T.,  
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T., Sano S.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RT "Complete sequencing and characterization of 21,243 full-length human  
CDNAs";  
RT Nat. Genet. 36:40-45 (2004).  
DR EMBL; AK094159; BAC04297.1; -.  
SQ SEQUENCE 126 AA; 14003 MW; AFI0B5375A3D9C7E CRC64;

Query March 57.5%; Score 42; DB 2; Length 126;  
Best Local Similarity 58.3%; Pred. No. 25;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLARA 14  
|||:|:|  
Db 108 GPDLRWAGSRA 119

RESULT 32  
Q8XP09 PRELIMINARY; PRT; 252 AA.  
AC Q8XP09;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.  
GN Name=RS02135; OrderedLocNames=RS021579;  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC plasmid megaplasmid.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Broctier P., Camus J.C., Cattolico L.,  
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502 (2002).  
CC -1- SIMILARITY: Contains 1 HTH LuxR-type DNA-binding domain.  
DR EMBL; AL646085; CAD18730.1; -.  
DR HSRP; P11470; 1FSE.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR Pfam; PF00196; GerP.1  
DR PRINTS; PR01590; HTHPIS.  
DR PRINTS; PR00038; HTHLXR.  
DR PRODOM; PD000307; HTH LuxR.1.  
DR SMART; SM00421; HTH LuxR.1.  
KW Complete proteome; DNA-binding; Plasmid; Transcription;  
KW Transcription regulation.  
SQ SEQUENCE 252 AA; 27666 MW; 483403EE326F7C2E CRC64;

Query March 57.5%; Score 42; DB 2; Length 252;  
Best Local Similarity 53.8%; Pred. No. 52;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAR 13  
|||:|:|  
Db 76 IDPTLRQWLATR 88

RESULT 33  
P90433 PRELIMINARY; PRT; 313 AA.  
ID P90433;  
AC P90433;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Truncated reverse transcriptase (fragment).  
GN Name=pol;  
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RX [1]  
RP SEQUENCE FROM N.A.  
RA Smith J.M., Krauselburd E.N., Torres J.V.;  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to peptidase family A2.  
DR EMBL; U83413; AAB41428.1; -.  
DR HSRP; Q07387; ITCW.  
DR MEROPS; A02.002; -.  
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
DR InterPro; IPR001995; Peptidase A2.  
DR InterPro; IPR009007; Peptidase A2.  
DR InterPro; IPR001969; Pept\_Asp\_AS.  
DR InterPro; IPR000477; RVTse.  
DR Pfam; PF00077; RVP.1.  
DR Pfam; PF00078; RVT.1.  
DR PROSITE; PS00141; ASP\_PROT\_RETROV.1.  
DR PROSITE; PS01175; ASP\_PROT\_RETROV.1.  
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;  
KW Transferase.  
FT NON TER 1  
FT NON TER 1  
SQ SEQUENCE 313 AA; 34674 MW; 5A0BB016783FC8A6 CRC64;

Query March 57.5%; Score 42; DB 2; Length 313;  
Best Local Similarity 87.5%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9  
|||:|:|  
Db 184 EGPTLRQW 191

RESULT 34  
O855N9 PRELIMINARY; PRT; 325 AA.  
ID O855N9;  
AC O855N9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Gp74.
OC Mycobacteriophage Cheqd.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
ON NCBI_TaxId=205876;
RX SEQUENCE FROM N.A.
RX MEDLINE=2259260; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Madeworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandassamy J., Keenan L., Bardarov S.,
RA Krizhov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL: AY129336; AAC07992.1; -.
SQ SEQUENCE 325 AA; 35999 MW; 04265796D0B4FC1D CRC64;

Query Match 57.5%; Score 42; DB 2; Length 325;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAR 14
Db 302 DGPTVQELAR 314

RESULT 35
P96I3 PRELIMINARY; PRT; 326 AA.
ID P96I3;
AC P96I3;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, last sequence update)
DR 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxId=399;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMB1;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstrom K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
DR EMBL: Y08963; CAJ70157.1; -.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR000847; HTH_LysR.
DR InterPro: IPR005119; LysR_subst.
DR InterPro: IPR009058; Wng_hlx_DNA_bnd.
DR Pfam: PF00126; HTH_1; 1.
DR Pfam: PF03466; LysR_substrate; 1.
DR PRINTS: PR00039; HTHLYSR.
DR PROSITE: PSS0931; HTH_LysR; 1.
KM DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BPE9C32F6719E28B CRC64;

Query Match 57.5%; Score 42; DB 2; Length 326;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAR 13
Db 204 KGPBLQWLSSQ 215

RESULT 36
Q7XPP6 PRELIMINARY; PRT; 375 AA.
ID Q7XPP6
AC Q7XPP6;

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DE OSJNB0053K19.27 protein (OSJNB0060E08.2 protein).
GN Name=OSJNB0053K19.27; Synonyms=OSJNB0060E08.2;
OS Oryza sativa [aponica cultivar-group].
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriocarideae; Oryzae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
RP SEQUENCE FROM N.A.
RP PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL: AL606645; CAB03519.2; -.
DR EMBL: AL606669; CAB04739.1; -.
DR Gramene; Q7XPP6; -.
DR GO: GO:0005515; F:protein binding; IEA.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR002083; MATF.
DR InterPro: IPR008974; Traf_like.
DR Pfam: PF00651; BTB; 1.
DR Pfam: PF00917; MATF; 1.
DR SMART: SM00225; BTB; 1.
DR SMART: SM00061; MATF; 1.
DR PROSITE: PSS0097; BTB; 1.
DR PROSITE: PSS0144; MATF; 1.
SQ SEQUENCE 375 AA; 41043 MW; 20FC6B99E4750816 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAR 13
Db 138 MERPRRLQWLRR 150

RESULT 37
Q9SLB9 PRELIMINARY; PRT; 450 AA.
ID Q9SLB9
AC Q9SLB9; Q9ASU2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DE Expressed protein (At2g42400/MRK10.12) (transcription factor
DE AtVO22).
GN Name=At2g42400; Synonyms=AtVO22;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Koning C.M., Benito M.-I.,
RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Nieman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee C.J., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Tanizaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shin P., Banh J., Bowser L.,
RA Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee C.J.,
RA Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
RA Southwick A., Toriumi M., Yamada K., Yu G., Shinzaki K., Davis R.W.,
RA Theologis A., Ecker J.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Mitsuda N., Hisabori T., Takeyasu K., Sato M.H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005956; AAD23723.2; -
DR EMBL; AY078048; AAL77749.1; -
DR EMBL; AF61831; AAK32843.1; -
DR EMBL; AB125257; BAD17858.1; -
DR PTR; E84853; E84853;
DR INTERPRET: IPR009105; Colicin_E3_cat.
SQ SEQUENCE 450 AA; 5056 MW; 44DBE7B4B69B95 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 450;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGETLRQWL 10
:|:|:|:|:|
Db 320 VEGETIREWL 329

RESULT 38
Q9N6P9 PRELIMINARY; PRT; 586 AA.
ID Q9N6P9;
AC Q9N6P9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE L354.2;
GN Name=L354.2;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fredlin;
RA Myler P.J., Sisk E., Cawthra J., Handley F., Vogt C., Robertson L.,
RA McDougall P., Ivens A., Nguyen D., Munden H., Stuart K.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC022473; AAF69566.1; -
SQ SEQUENCE 586 AA; 6394 MW; 208367C213896AF3 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 586;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IEGETLRQWL 14
:|:|:|:|:|
Db 58 VEAPLITQMTMAA 71

```

```

RESULT 39
POL_SIVS4 STANDARD; PRT; 1019 AA.
ID POL_SIVS4
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Pol polyprotein (Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)).
GN Name=POL;
OS Simian immunodeficiency virus (P236/smH4 isolate) (sooty mangabey).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99262053; Pubmed=2786147; DOI=10.1038/339389a0;
RA Hirsch V.M., Olmstead R.A., Murphy-Cord M., Purcell R.H.,
RA Johnson P.R.;
RL "An African primate lentivirus (SIVsm) closely related to HIV-2."
RL Nature 339:389-392(1989)
CC -1- FUNCTION: During replicative cycle of retroviruses, the reverse-
CC transcribed viral DNA is integrated into the host chromosome by
CC the viral integrase enzyme. RNase H activity is associated with
CC the reverse transcriptase.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(n).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC -1- SIMILARITY: Contains 1 integrase-type zinc finger.
CC -1- SIMILARITY: Contains 1 peptidase A2 domain.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC -1- SIMILARITY: Contains 1 RNase H domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14307; -; NOT_ANNOTATED_CDS.
DR HSSP; P04584; IMU2.
DR MEROPS; A02.002; -.
DR HTY; X14307; POL:SSMWH4.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn_N.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_AspArtic.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR004777; RVTse.
DR InterPro; IPR004777; RVTse.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_chumb.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_1.
DR Pfam; PF00075; RNaseH_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR Pfam; PF06815; RVT_chumb; 1.
DR Pfam; PF06817; RVT_chumb; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50878; RNASE_H; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR AIDS; Aspartyl protease; DNA integration; DNA recombination;

```

KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;  
KM Nuclease; Polypeptide; RNA-directed DNA polymerase; Transferase; Zinc;  
KM Zinc-finger.  
FT CHAIN 1 167 Protease.  
FT DOMAIN 211 401 Reverse transcriptase.  
FT DOMAIN 600 723 RNase H.  
FT ZN\_FING 729 770 Integrase-type.  
FT ACT\_SITE 93 93 By similarity.  
SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92BLC CRC64;  
Query Match 57.5%; Score 42; DB 1; Length 1019;  
Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 EGPTRLROW 9  
Db 184 EGPTRLROW 191  
RESULT 40  
P89154 PRELIMINARY; PRT; 1019 AA.  
AC P89154;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pol polyprotein (Fragment).  
GN Name=pol;  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN SEQUENCE FROM N.A.  
RC STRAIN=STVsm543;  
RX MEDLINE=97151152; PubMed=8995688;  
RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,  
RA Elkins W.R., Montefiori D.C.;  
RT "A molecularly cloned, pathogenic, neutralization-resistant simian  
RT immunodeficiency virus, STVsm543-3.";  
RL J. Virol. 71:1608-1620 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=STVsm543;  
RA Hirsch V.M.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family A2.  
DR EMBL: U72748; AAC56559.1; -.  
DR PIR: T11560; T11560.  
DR HSSP: P04584; 1MU2.  
DR MEROP; A02.002; -.  
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
DR GO: GO:0003677; F:DNA binding; IEA.  
DR GO: GO:0008907; F:integrase activity; IEA.  
DR GO: GO:0008233; F:peptidase activity; IEA.  
DR GO: GO:0008233; F:ribonuclease H activity; IEA.  
DR GO: GO:0003723; F:RNA binding; IEA.  
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
DR GO: GO:0016740; F:transferase activity; IEA.  
DR GO: GO:0008270; F:zinc ion binding; IEA.  
DR GO: GO:0015074; F:DNA integration; IEA.  
DR GO: GO:0006310; P:DNA recombination; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_Zn\_N.  
DR InterPro: IPR001995; Peptidase\_A2.  
DR InterPro: IPR009007; Pept\_Aspartic.  
DR InterPro: IPR001969; Pept\_Asp\_AS.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; Rve.  
DR InterPro: IPR000477; RVTse.  
DR InterPro: IPR010659; RVT\_connect.  
DR InterPro: IPR010661; RVT\_thumb.  
DR InterPro: IPR010661; RVT\_thumb.

DR InterPro: IPR005829; Sug\_transporter.  
DR Pfam: PF02022; Integrase\_Zn\_1.  
DR Pfam: PF00075; RNaseH\_1.  
DR Pfam: PF00665; rve\_1.  
DR Pfam: PF00078; RVP\_1.  
DR Pfam: PF00078; RVP\_1; 1.  
DR Pfam: PF06815; RVT\_connect; 1.  
DR Pfam: PF06817; RVT\_thumb\_1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
DR PROSITE: PS00217; SUGAR\_TRANSPORT\_2; UNKNOWN\_1.  
KM Aspartyl protease; Hydrolase; Polypeptide; Protease;  
KM RNA-directed DNA polymerase; Transferase.  
FT NON\_TER 1 1  
SQ SEQUENCE 1019 AA; 115595 MW; 26F1EF4594E59537 CRC64;  
Query Match 57.5%; Score 42; DB 2; Length 1019;  
Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 EGPTRLROW 9  
Db 184 EGPTRLROW 191  
RESULT 41  
Q7ZBR5 PRELIMINARY; PRT; 1019 AA.  
AC Q7ZBR5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Pol (fragment).  
GN Name=pol;  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22628501; PubMed=12743298;  
RX DOI=10.1128/JVI.77.11.6405-6418.2003;  
RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;  
RT "Unique pattern of convergent envelope evolution in simian  
RT immunodeficiency virus-infected rapid progressor macaques: association  
RT with CD4-independent usage of CCR5.";  
RL J. Virol. 77:6405-6418 (2003).  
CC -!- SIMILARITY: Belongs to peptidase family A2.  
DR EMBL: AY221515; AAC67309.1; -.  
DR HSSP: P04584; 1MU2.  
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
DR GO: GO:0003677; F:DNA binding; IEA.  
DR GO: GO:0008907; F:integrase activity; IEA.  
DR GO: GO:0008233; F:peptidase activity; IEA.  
DR GO: GO:0004523; F:ribonuclease H activity; IEA.  
DR GO: GO:0003723; F:RNA binding; IEA.  
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
DR GO: GO:0016740; F:transferase activity; IEA.  
DR GO: GO:0008270; F:zinc ion binding; IEA.  
DR GO: GO:0015074; F:DNA integration; IEA.  
DR GO: GO:0006310; P:DNA recombination; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_Zn\_N.  
DR InterPro: IPR001995; Peptidase\_A2.  
DR InterPro: IPR009007; Pept\_Aspartic.  
DR InterPro: IPR001969; Pept\_Asp\_AS.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; Rve.  
DR InterPro: IPR000477; RVTse.  
DR InterPro: IPR010659; RVT\_connect.  
DR InterPro: IPR010661; RVT\_thumb.  
DR InterPro: IPR005829; Sug\_transporter.

DR Pfam; PF02022; Integrase\_Zn; 1.  
 DR Pfam; PF00075; RNaseH; 1.  
 DR Pfam; PF00665; rve; 1.  
 DR Pfam; PF00077; RVP; 1.  
 DR Pfam; PF00078; RVP; 1.  
 DR Pfam; PF06817; RVT\_connect; 1.  
 DR Pfam; PF06817; RVT\_chumb; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE; PS50175; ASP\_PROT\_RETROV; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; UNKNOWN 1.  
 KM Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;  
 KW Transferrase.  
 FT NON\_TER  
 SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;  
 Query Match 57.5%; Score 42; DB 2; Length 1019;  
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9  
 Db 184 EGPKLRQW 191  
 RESULT 42  
 072BR7 PRELIMINARY; PRT; 1019 AA.  
 AC 072BR7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Pol (Fragment).  
 GN Name=pol;  
 OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirinae.  
 OC NCBI\_TaxID=11723;  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22628501; PubMed=12743298;  
 RX DOI=10.1126/JVI.77.11.6405-6418.2003;  
 RA Deighan H., Puffer B.A., Dome R.W., Hirsch V.M.;  
 RT Unique pattern of convergent envelope evolution in simian  
 RT immunodeficiency virus-infected rapid progressor macaques: association  
 RT with CD4-independent usage of CCR5.";  
 RL J. Virol. 77:6405-6418(2003).  
 CC -1- SIMILARITY: Belongs to peptidase family A2.  
 DR EMBL; AY221514; AA067307.1; -.  
 DR HSPSP; P04584; IIMU2.  
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0008907; F:integrase activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0016740; F:transferrase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0015074; P:DNA recombination; IEA.  
 DR GO; GO:0006310; P:DNA integration; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
 DR InterPro; IPR001037; Integrase\_C.  
 DR InterPro; IPR003308; Integrase\_Zn\_N.  
 DR InterPro; IPR001995; Peptidase\_A2.  
 DR InterPro; IPR009007; Pept\_Asp\_AS.  
 DR InterPro; IPR001969; Pept\_Asp\_AS.  
 DR InterPro; IPR001561; RNaseH.  
 DR InterPro; IPR001584; Rve.  
 DR InterPro; IPR000477; RVTse.  
 DR InterPro; IPR010659; RVT\_connect.  
 DR InterPro; IPR010661; RVT\_chumb.  
 DR InterPro; IPR005829; Sug\_transporter.  
 DR Pfam; PF02022; Integrase\_Zn; 1.

DR Pfam; PF00075; RNaseH; 1.  
 DR Pfam; PF00665; rve; 1.  
 DR Pfam; PF00077; RVP; 1.  
 DR Pfam; PF00078; RVP; 1.  
 DR Pfam; PF06817; RVT\_connect; 1.  
 DR Pfam; PF06817; RVT\_chumb; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE; PS50175; ASP\_PROT\_RETROV; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; UNKNOWN 1.  
 KM Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;  
 KW Transferrase.  
 FT NON\_TER  
 SQ SEQUENCE 1019 AA; 115340 MW; A88652D5F1BE26F CRC64;  
 Query Match 57.5%; Score 42; DB 2; Length 1019;  
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9  
 Db 184 EGPKLRQW 191  
 RESULT 43  
 08P9L5 PRELIMINARY; PRT; 410 AA.  
 AC 08P9L5;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Valine-pyruvate aminotransferase.  
 GN Name=avta; OrderedLocNames=XCC1839;  
 OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OC NCBI\_TaxID=340;  
 OX NCBI\_TaxID=340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 33913 / NCPPB 528;  
 RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolin L.M.C., Canargo L.E.A.,  
 RA Catarote G., Camavan F., Cardozo J., Chambergo F., Clapina L.P.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Medeiros J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Martins E.C., Meidanis J., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Secubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AE012286; AA41127.1; -.  
 DR GO; GO:0008483; F:transaminase activity; IEA.  
 DR GO; GO:0016740; F:biosynthesis; IEA.  
 DR GO; GO:0009058; F:diolysis; IEA.  
 DR InterPro; IPR004839; Aminotrans\_1/II.  
 DR Pfam; PF00155; Aminotran\_1\_2; 1.  
 KM Aminotransferase; Complete proteome; Pyruvate; Transferrase.  
 SQ SEQUENCE 410 AA; 44530 MW; 441BEA9A9F04F553 CRC64;  
 Query Match 56.8%; Score 41.5; DB 2; Length 410;  
 Best Local Similarity 56.2%; Pred. No. 1e+02;  
 Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Db 78 GPTGYAPLREWVAAR 93

## RESULT 44

Q9PLE2 PRELIMINARY; PRT; 427 AA.

AC Q9PLE2; 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Value-private aminotransferase.  
 GN Name-avfA, OrderedlocusNames=XAC1858;  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Furlan L.R.,  
 RA Ouaggo R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Chapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities."  
 RL Nature 417:459-463(2002).  
 DR EMBL; AE011819; AAM36720.1; -.  
 DR GO; GO:0008483; F:transaminase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR004839; AminoTrans\_I/II.  
 DR Pfam; PF00155; AminoTrans\_1\_2; 1.  
 DR Complete proteome.  
 SK SEQUENCE 427 AA; 45926 MW; BBD205BDA06E56E CRC64;

Query Match 56.8%; Score 41.5; DB 2; Length 427;

Best Local Similarity 56.2%; Pred. No. 1.1e+02; Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy 3 GPT-----LRQWLAR 13  
 ||| |||  
 Db 95 GPTGYAPLREWVAAR 110

## RESULT 45

Q98AJ1 PRELIMINARY; PRT; 75 AA.

AC Q98AJ1; 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, last annotation update)  
 DE Transposase.  
 GN OrderedlocusNames=msr5979;  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Ideesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti."  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003008; BAB52338.1; -.  
 KW Complete proteome.  
 SK SEQUENCE 75 AA; 8363 MW; B76547C20DA52E4D CRC64;

Query Match 56.2%; Score 41; DB 2; Length 75;

Best Local Similarity 58.3%; Pred. No. 22; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAR 13  
 :| | | | |  
 Db 3 QGKACREWLAR 14

Search completed: September 1, 2005, 16:21:19  
 Job time : 55.0719 secs